

Rooke A.  
101762588 Page 1  
Seq ID 16Gencore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2006, 13:07:03 ; Search time 199 Seconds  
(without alignments)  
35.327 Million cell updates/secTitle: US-10-762-588-1  
Sequence: 1 HIIHNVKKEHAAHN 16Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5Searched: 2443163 seqs, 439378781 residues  
Total number of hits satisfying chosen parameters: 2443163Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database :

1: Genesegp1980s:\*  
2: genesegp1990s:\*  
3: genesegp2000s:\*  
4: genesegp2001s:\*  
5: genesegp2002s:\*  
6: genesegp2003as:\*  
7: genesegp2003bs:\*  
8: genesegp2004s:\*  
9: genesegp2005s:\*Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	100.0	16	6	ABP97774 Amino act
2	95	100.0	16	6	ABU08453 Affinity
3	95	100.0	16	6	ABU08446 Affinity
4	95	100.0	16	6	ADR70438 Metal ion
5	95	100.0	20	7	ADP55558 Histidine
6	95	100.0	20	7	ADP56001 His tag p
7	95	100.0	20	7	ADP59859 C-termina
8	95	100.0	20	9	ADZ02850 Baker's y
9	95	100.0	32	3	AAV44208 Fusion pr
10	95	100.0	45	3	AAV44208 Protein c
11	95	100.0	211	3	ADZ67254 Escherich
12	95	100.0	278	3	AAV44207 Affinity
13	95	100.0	340	6	ABU08445 Recombina
14	95	100.0	19	7	ADZ6604 Histidine
15	87	81.1	19	7	ADG14415 Chicken 1
16	57	60.0	69	5	ABP03327 Human ORF
17	55	55.8	837	4	ABP03327 Human ORF
18	51	53.7	402	4	ABP03327 Human ORF
19	50	52.6	299	3	AAV44207 Affinity
20	50	52.6	300	3	AAV44207 Affinity
21	50	52.6	351	2	AAV44207 Affinity
22	49.5	52.1	198	9	ADM13460 E. tenell
23	49.5	52.1	334	8	ADM13460 E. tenell
24	49.5	52.1	334	8	ADM13460 E. tenell

25	49	51.6	117	4	AAO10111 Human pol
26	49	51.6	134	3	AAO10111 Human pol
27	49	51.6	368	4	ABP69867 Drosophila
28	49	51.6	392	8	ADT59295 Plant pol
29	49	51.6	418	8	ADP68442 Plant pol
30	49	51.6	459	8	AAO17798 Rice leaf
31	49	51.6	459	7	ABP67520 Rice leaf
32	49	51.6	506	7	ADM17713 Plant rad
33	49	51.6	1463	2	AAW9482 Murine NC
34	48	50.5	74	5	ABP07487 Human ORF
35	48	50.5	86	8	ADN73121 Thale cre
36	48	50.5	161	7	ABP62756 Klebsiell
37	48	50.5	191	8	ADP17428 Amino act
38	48	50.5	400	7	ADP05866 Bacterial
39	48	50.5	401	4	AAE02505 Arabidops
40	48	50.5	401	7	ABO43132 A. thalia
41	48	50.5	401	7	AAE26991 Streses to
42	48	50.5	427	8	ADP01779 Thalecres
43	48	50.5	466	5	AAU92998 Arabidops
44	48	50.5	466	7	ADP05993 Plant yie
45	48	50.5	466	7	ADP1543 Plant yie
46	48	50.5	466	8	ADP13817 Plant tra
47	48	50.5	470	8	ADP72429 Plant ful
48	48	50.5	484	9	ADM17335 Eucalyptu
49	48	50.5	519	6	ADA54404 Human pro
50	48	50.5	670	7	ADM04418 Human pro
51	48	50.5	684	5	ABP62691 Drosophila
52	48	50.5	684	5	AAU74625 Oestrogen
53	47	49.5	124	4	ABG12879 Novel hum
54	47	49.5	265	8	ADY06478 Plant ful
55	47	49.5	270	8	ADY22561 Plant ful
56	47	49.5	470	4	ABP61018 Drosophila
57	47	49.5	517	4	ABP61245 Drosophila
58	47	49.5	639	8	ADP99031 C. albica
59	47	49.5	645	4	ABP61545 Drosophila
60	47	49.5	697	4	ABP68377 Drosophila
61	46.5	48.9	352	9	ABP39742 L. pneumo
62	46.5	48.9	356	9	ABP36325 L. pneumo
63	46	48.4	47	3	AAV93345 Light cha
64	46	48.4	58	4	AAV93345 Light cha
65	46	48.4	62	3	AAV93345 Light cha
66	46	48.4	63	2	AAV93345 Light cha
67	46	48.4	77	5	ABP01829 Human ORF
68	46	48.4	83	3	AAV93347 Light cha
69	46	48.4	94	3	AAV93351 Light cha
70	46	48.4	109	8	ADP93148 Novel hum
71	46	48.4	109	8	ADP93188 Novel hum
72	46	48.4	125	5	ABP78708 Human hig
73	46	48.4	125	8	ADH10414 Human kin
74	46	48.4	126	8	ADP72278 Plant ful
75	46	48.4	131	2	AAV75181 Partial p
76	46	48.4	142	4	AAO10180 Human pol
77	46	48.4	158	5	ABP78709 Calmoduli
78	46	48.4	179	3	AAV93353 Light cha
79	46	48.4	186	3	AAV93349 Light cha
80	46	48.4	243	8	ADP93190 Novel hum
81	46	48.4	243	8	ADP93152 Novel hum
82	46	48.4	255	4	AAV93342 Light cha
83	46	48.4	255	4	AAV93342 Light cha
84	46	48.4	266	9	AEA21851 Human kin
85	46	48.4	272	9	AEA21851 Human kin
86	46	48.4	277	9	ABP70078 Drosophila
87	46	48.4	292	6	ABU42182 Protein e
88	46	48.4	305	4	ABG21100 Novel hum
89	46	48.4	320	5	ADP97644 Scaphyloc
90	46	48.4	330	6	ABP17186 Scaphyloc
91	46	48.4	357	6	ABP17186 Scaphyloc
92	46	48.4	363	6	ABU21590 Protein e
93	46	48.4	415	6	ABU21590 Protein e
94	46	48.4	415	8	ADM93857 Human NOV
95	46	48.4	415	8	ADM93166 Human NOV
96	46	48.4	579	9	ABP32342 Human pro
97	46	48.4	579	9	ABP32338 Human pro

98	46	48.4	589	4	ABB58929	Abb58929	Drosophila
99	46	48.4	612	8	ADS93142	Ads93142	Novel hum
100	46	48.4	615	6	ABU9144	Abu9144	Novel hum

## ALIGNMENTS

```

RESULT 1
ABP97774
ID ABP97774 standard; peptide; 16 AA.
XX
XX ABP97774;
AC
XX
DT 28-MAY-2003 (first entry)
XX
DE Amino acid sequence of an affinity peptide.
XX
XX Polymeric metal ion affinity compound; protein purification;
KW aspartate based metal chelating ligand; metal ion; analyte detection;
XX water-soluble polymeric substrate; affinity peptide.
XX
XX Unidentified.
OS
XX
XX WO200300708-A1.
XX
XX 03-JAN-2003.
XX
XX 20-JUN-2002; 2002MO-US019879.
XX
XX 21-JUN-2001; 2001US-0300336P.
XX
XX (CLON-) CLONTECH LAB INC.
XX
XX Tchnaga GS; ✓
XX
XX WPI; 2003-247890/24.
XX
XX Water-soluble polymeric metal ion affinity compounds useful in e.g.
PT analyte detection, comprises an aspartate based metal chelating ligand
PT bonded to a water-soluble polymeric substrate.
XX
XX Disclosure; Page 15; 36pp; English.
XX
XX The specification describes a water-soluble polymeric metal ion affinity
XX compound, comprising an aspartate based metal chelating ligand bonded to
XX a water-soluble polymeric substrate. The ligand is complexed with a metal
XX ion. The water soluble metal ion affinity reagents have additional
XX selectivity for protein purification applications and novel analyte
XX detection applications. The compounds of the invention are useful in a
XX variety of different applications, including analyte detection and
XX purification applications. The present sequence represents an affinity
XX peptide, which is part of an analyte which can be purified using the
XX compounds of the invention
XX
XX Sequence 16 AA;
SQ
Query Match 100.0%; Score 95; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
OY 1 HLINNVAKKEHHAAHN 16
|||
|||
|||
DB 1 HLINNVAKKEHHAAHN 16
|||
|||
|||
RESULT 2
ABU08453
ID ABU08453 standard; peptide; 16 AA.
XX
XX ABU08453;
AC
XX
DT 17-JUN-2003 (first entry)

```

XX Affinity peptide #2.

XX  
XX  
XX Metal ion affinity peptide; fusion protein; protein purification; Ca2+;  
XX metal chelating resin; metal ion chelate resin; hard metal; Fe3+;  
XX Al3+; Co2+; Cu2+; Ni2+; Zn2+; Co2+; protein-protein interaction;  
XX metal ion affinity peptide-tagged recombinant protein; IMAC;  
XX DNA-protein interaction; immobilised metal ion affinity chromatography;  
XX gene expression; phosphorylation state.

XX  
XX  
XX Synthetic.

XX  
XX  
XX US2002164718-A1.

XX  
XX  
XX 07-NOV-2002.

XX  
XX  
XX 15-MAY-2001; 2001US-00858332.

XX  
XX  
XX 25-SEP-1998; 98US-0101867P.  
XX 23-SEP-1999; 99US-00404017.

XX  
XX  
XX (TCHG/) TCHGA G S.  
XX (JOKH/) JOKHADZE G G.

XX  
XX  
XX Tchaga GS, Jokhadze GG;

XX  
XX  
XX WPI; 2003-361747/34.

XX  
XX  
XX New metal ion affinity peptide useful, when fused to a fusion partner  
XX polypeptide, for protein purification methods and to study protein-  
XX protein interactions and nucleic acid-protein interactions.

XX  
XX  
XX Claim 8; Page 13; 23pp; English.

XX  
XX  
XX The present invention relates to metal ion affinity peptides, fusion  
XX proteins containing metal ion affinity peptides, and polynucleotide  
XX sequences encoding the fusion proteins. The presence of a metal ion  
XX affinity peptide in a fusion protein allows purification of the fusion  
XX protein on a metal chelating resin. The method involves contacting a  
XX sample comprising a fusion protein with a metal ion chelate resin  
XX comprising a first metal ion, preferably a hard metal ion such as Fe3+,  
XX Ca2+ or Al3+ and eluting any resultant bound fusion protein from the  
XX resin. The resin comprises an immobilised Co2+ ion. The method further  
XX comprises contacting the sample with a second immobilised metal ion  
XX affinity resin comprising a second immobilised metal ion and eluting any  
XX resultant bound fusion protein from the first and second resins. The  
XX second metal ion is an intermediate metal ion such as Cu2+, Ni2+ or  
XX Co2+. The metal ion affinity peptide-tagged recombinant proteins are  
XX useful for the study of protein-protein interactions and nucleic acid  
XX molecule-protein interactions, using solid phase immobilised metal ion  
XX affinity chromatography (IMAC). They are also useful in high throughput  
XX systems which find use in massive parallel gene expression experiments,  
XX e.g. to determine the effect of an agent on synthesis of a protein or set  
XX of proteins, to analyse developmental stage-specific or tissue-specific  
XX synthesis of a protein and to analyse the phosphorylation state of a  
XX protein. These methods find use in applications to characterise a protein  
XX of unknown identity or function, and in enzymatic reactions. The present  
XX sequence represents an affinity peptide

XX  
XX  
XX Sequence 16 AA:

XX  
XX  
XX Query Match 100.0%; Score 95; DB 6; Length 16;  
XX Best Local Similarity 100.0%; Pred. No. 3.9e-08;  
XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX  
XX  
XX 1 HLINVHKERAAAHN 16  
XX |||||||||||||  
XX 1 HLINVHKERAAAHN 16

XX  
XX  
XX RESULT 3  
XX ABU08446  
XX ID ABU08446 standard; peptide; 16 AA.

ABU08446;  
17-JUN-2003 (first entry)  
Affinity purification peptide #1.

Metal ion affinity peptide; fusion protein; protein purification;  
metal chelating resin; metal ion chelate resin; hard metal; Fe<sup>3+</sup>; Ca<sup>2+</sup>;  
Al<sup>3+</sup>; Co<sup>2+</sup>; Cu<sup>2+</sup>; Ni<sup>2+</sup>; Zn<sup>2+</sup>; Co<sup>2+</sup>; protein-protein interaction;  
metal ion affinity peptide-tagged recombinant protein; IMAC;  
DNA-protein interaction; immobilised metal ion affinity chromatography;  
gene expression; phosphorylation state.

Synthetic.

US2002164718-A1.

07-NOV-2002.

15-MAY-2001; 2001US-00858332.

25-SEP-1998; b8US-0101867P.  
23-SEP-1999; 99US-00404017.

(TCHAG/) TCHAGA G S. ✓  
(JOKH/) JOKHADZE G G.

Tchaga GS, Jokhadze GG;  
WPI: 2003-361747/34.  
DR N-PSDB; ABX94279.

New metal ion affinity peptide useful, when fused to a fusion partner polypeptide, for protein purification methods and to study protein-protein interactions and nucleic acid-protein interactions.

Example 1; Fig 3; 23pp; English.

The present invention relates to metal ion affinity peptides, fusion proteins containing metal ion affinity peptides, and polynucleotide sequences encoding the fusion proteins. The presence of a metal ion affinity peptide in a fusion protein allows purification of the fusion protein on a metal chelating resin. The method involves contacting a sample comprising a fusion protein with a metal ion chelate resin comprising a first metal ion, preferably a hard metal ion such as Fe<sup>3+</sup>, Ca<sup>2+</sup> or Al<sup>3+</sup> and eluting any resultant bound fusion protein from the resin. The resin comprises an immobilised Co<sup>2+</sup> ion. The method further comprises contacting the sample with a second immobilised metal ion affinity resin comprising a second immobilised metal ion and eluting any resultant bound fusion protein from the first and second resins. The second metal ion is an intermediate metal ion such as Cu<sup>2+</sup>, Ni<sup>2+</sup>, Zn<sup>2+</sup> or Co<sup>2+</sup>. The metal ion affinity peptide-tagged recombinant proteins are useful for the study of protein-protein interactions and nucleic acid molecule-protein interactions, using solid phase immobilised metal ion affinity chromatography (IMAC). They are also useful in high throughput systems which find use in massive parallel gene expression experiments, e.g. to determine the effect of an agent on synthesis of a protein or set of proteins, to analyse developmental stage-specific, or tissue-specific synthesis of a protein and to analyse the phosphorylation state of a protein. These methods find use in applications to characterise a protein of unknown identity or function, and in enzymatic reactions. ABU08446-ABU08450 represent affinity purification peptides

Sequence 16 AA;

Query Match 100.0%; Score 95; DB 6; Length 16;  
Best Local Similarity 100.0%; Pred. No. 3,9e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 HLINHVKEEHAHAHN 16  
|||||  
1 HLINHVKEEHAHAHN 16

```

RESULT 5
ADR70438
ID ADR70438 standard; peptide; 16 AA.
XX
XX
AC ADR70438;
XX
XX
DT 02-DEC-2004 (first entry)
XX
XX
DE Metal ion affinity peptide seqd 1.
XX
XX
KW fusion protein; metal ion affinity peptide; metal ion chelate resin;
XX
XX
KW immobilised metal ion; purification.
XX
XX
OS Unidentified.
XX
XX
US US2004180415-A1.
XX
XX
PN 16-SEP-2004.
XX
XX
PD 21-JAN-2004; 2004US-00762588.
XX
XX
PF 15-MAY-2001; 2001US-00858332.
XX
XX
PR 21-JAN-2003; 2003US-0441804P.
XX
XX
PA (TCHG//) TCHAG S.
XX
XX
PA (JOKH//) JOKHADZE G. G.
XX
XX
PI Tchaga GS, Jokhadze GG;
XX
XX
DR WPI; 2004-675606/66.
XX
XX
PT Purification of fusion protein comprising metal ion affinity peptide, by
PT contacting sample comprising fusion protein with different metal ion
PT chelate resins with respective immobilized metal ion.
XX
XX
PS Disclosure; SEQ ID NO 1; 20pp; English.
XX
XX
CC The invention describes a fusion protein purified by contacting a sample
CC comprising fusion protein comprising a metal ion affinity peptide with
CC first metal ion chelate resin eluting bound fusion protein to produce a
CC first eluate; contacting the first eluate with a second metal ion
CC affinity resin; and eluting bound fusion protein to produce a product
CC eluate comprising a purified fusion protein. The chelate resins
CC respectively comprise immobilised metal ion. A fusion protein is purified
CC by contacting a sample comprising fusion protein having metal ion
CC affinity peptide with first metal ion chelate resin; eluting bound fusion
CC protein from the resin to produce a first eluate; contacting the first
CC eluate with a second metal ion affinity resin; and eluting bound fusion
CC protein from the resins to produce a product eluate comprising a purified
CC fusion protein. The chelate resins respectively comprise immobilised
CC metal ion. Also described is a kit for purifying protein comprising first
CC and second metal ion chelate resins with respective immobilised metal
CC ion. The method is useful for purifying a fusion protein having metal ion
CC affinity peptide. The use of two different metal ions for purification of
CC protein tagged with a single metal ion affinity peptide provides high
CC degree of purification with a single chromatographic step. This is the
CC amino acid sequence of a metal ion affinity peptide that can be used in
CC the creation of fusion proteins of the invention.
XX
XX
SQ Sequence 16 AA;
XX
XX
Query Match 100.0%; Score 95; DB 8; Length 16;
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 HLINHVKEEHAHAN 16
Db 1 HLINHVKEEHAHAN 16

```

ADFS5558  
 ID ADFS5558 standard; peptide; 20 AA.  
 XX  
 AC ADFS5558;  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Histidine tag.  
 XX  
 KM Histidine tag; gene; ds; cell-free protein synthesis.  
 XX  
 OS Synthetic.  
 PN JP2003009880-A.  
 XX  
 PD 14-JAN-2003.  
 XX  
 PF 02-JUL-2001; 2001JP-00201356.  
 XX  
 PR 02-JUL-2001; 2001JP-00201356.  
 XX  
 PA (RIKA ) RIKAGAKU KENKYUSHO.  
 XX  
 DR WPI; 2003-590463/56.  
 XX  
 PT Manufacturing template DNA for cell-free protein synthesis comprises  
 PT amplifying a double-stranded DNA fragment which encodes protein, and DNA  
 PT fragments which overlap 5' and 3' ends of the coding sequence.  
 XX  
 PS Claim 12; SEQ ID NO 1; 38pp; Japanese.  
 XX  
 CC The invention relates to a method for manufacturing template DNA for  
 CC protein synthesis comprising PCR amplifying a first double-stranded DNA  
 CC (dsDNA) fragment which encodes a protein and contacting the DNA with  
 CC second and third dsDNA fragments which overlap the 5' and 3' ends of the  
 CC first DNA fragment respectively, where sense and antisense primers which  
 CC anneal to the 5' and 3' ends of the second and third fragments  
 CC respectively, are used to initiate PCR. The invention also relates to a  
 CC cell-free protein synthesis method which involves the use of a template  
 CC DNA produced by the method of the invention. PCR amplifying DNA  
 CC comprising primary and secondary PCR steps is useful for synthesizing a  
 CC template DNA for protein synthesis, preferably cell-free protein  
 CC synthesis. The template DNA (based on cDNA or genomic DNA sequences) for  
 CC expressing and purifying protein can be manufactured quickly and  
 CC efficiently. The structure and function relationship of a protein can be  
 CC analysed quickly. This sequence represents a histidine tag used in the  
 CC method of the invention.  
 CC  
 SQ Sequence 20 AA;  
 XX  
 Query Match 100.0%; Score 95; DB 7; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 5.1e-08;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 HLINHVHKEHAAHN 16  
 |||||||||  
 DB 4 HLINHVHKEHAAHN 19

RESULT 6  
 ADJ66001  
 ID ADJ66001 standard; peptide; 20 AA.  
 XX  
 AC ADJ66001;  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE His tag peptide, SEQ ID 1.  
 XX  
 KM PCR, DNA preparation; protein production; His tag.  
 XX  
 OS Synthetic.  
 XX

PN JP2003009877-A.  
 XX  
 PD 14-JAN-2003.  
 XX  
 PF 02-JUL-2001; 2001JP-00200676.  
 XX  
 PR 02-JUL-2001; 2001JP-00200676.  
 XX  
 PA (RIKA ) RIKAGAKU KENKYUSHO.  
 XX  
 DR WPI; 2003-601093/57.  
 XX  
 PT Efficient PCR-based method for preparation of a template DNA for  
 PT production of a protein using an acellular protein synthetic system.  
 XX  
 PS Claim 12; SEQ ID NO 1; 17pp; Japanese.  
 XX  
 CC The present invention relates to a PCR-based method for preparing a  
 CC template DNA for production of a protein using an acellular protein  
 CC synthetic system. The method comprises performing PCR on a reaction  
 CC mixture comprising: a first double-stranded DNA fragment comprising a  
 CC sequence encoding a protein or its fragment; a second double-stranded DNA  
 CC fragment comprising a sequence which overlaps with the 5'-terminal region  
 CC of the first DNA fragment; a third double-stranded DNA fragment which  
 CC overlaps with the 3'-terminal region of the first DNA fragment; a sense  
 CC primer annealing with the 5'-terminal region of the second DNA fragment;  
 CC and an antisense primer annealing with the 3'-terminal region of the  
 CC third DNA fragment. The present sequence was used to illustrate the  
 CC invention.  
 CC  
 SQ Sequence 20 AA;  
 XX  
 Query Match 100.0%; Score 95; DB 7; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 5.1e-08;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 HLINHVHKEHAAHN 16  
 |||||||||  
 DB 4 HLINHVHKEHAAHN 19

RESULT 7  
 ADZ39859  
 ID ADZ39859 standard; peptide; 20 AA.  
 XX  
 AC ADZ39859;  
 DT 30-JUN-2005 (first entry)  
 XX  
 DE C-terminal defective protein-related peptide - SEQ ID 11.  
 XX  
 KM protein engineering; protein solubilization.  
 XX  
 OS Unidentified.  
 XX  
 PN JP2005095079-A.  
 XX  
 PD 14-APR-2005.  
 XX  
 PF 25-SEP-2003; 2003JP-00333932.  
 XX  
 PR 25-SEP-2003; 2003JP-00333932.  
 XX  
 PA (DOKU-) DOKURITSU GYOSEI HOJIN RIKAGAKU KENKYUSHO.  
 XX  
 DR WPI; 2005-299480/31.  
 XX  
 PT Synthesizing C-terminal defective protein, involves translating 3'  
 PT terminal deleted RNA obtained by performing RNA transcription using  
 PT template DNA, RNA polymerase and substance capable of stopping RNA  
 PT transcription.  
 XX  
 PS Example 2; SEQ ID NO 11; 19pp; Japanese.  
 XX



XX The invention comprises a method of synthesizing a C-terminal defective  
CC protein, the method involves obtaining 3' terminal deleted RNA through  
CC RNA transcription, and obtaining C-terminal defective protein by  
CC translating the obtained 3' terminal deleted RNA. The method of the  
CC invention is useful for synthesizing C-terminal defective protein.  
CC preparing soluble protein, and for determining solubility domain. The  
CC present amino acid sequence represents a peptide that was used in an  
CC example of the invention.  
XX  
SQ Sequence 20 AA;  
Query Match 100.0%; Score 95; DB 9; Length 20;  
Best Local Similarity 100.0%; Pred. No. 5.1e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 HLHNHVKKEHHAAHN 16  
Db 4 HLHNHVKKEHHAAHN 19  
RESULT 8  
AD202850  
ID AD202850 standard; peptide; 20 AA.  
XX  
AC AD202850;  
XX  
DT 30-JUN-2005 (first entry)  
XX  
DE Baker's yeast YJL15W/protein production-related His tag peptide.  
XX  
KM protein production; protein engineering; polymorphism; drug metabolism.  
XX  
OS Synthetic.  
XX  
PN JP2005095083-A.  
XX  
PD 14-APR-2005.  
XX  
PF 25-SEP-2003; 2003JP-00334083.  
XX  
PR 25-SEP-2003; 2003JP-00334083.  
XX  
PA (DOKU-) DOKURITSU GYOSEI HOJIN RIKAGAKU KENKUTSUH.  
XX  
DR WPI; 2005-299482/31.  
XX  
PT Producing mutant protein by cell free protein synthesis system, comprises  
PT amplifying two double-stranded DNA fragments that introduce linker and  
PT mutation sequences, connecting DNA fragments and amplifying linear double  
PT -stranded DNA.  
XX  
PS Example 1; SEQ ID NO 27; 22pp; Japanese.  
XX  
CC The invention relates to a novel method for producing a variant or mutant  
CC protein using a cell-free protein synthesis system. The method comprises  
CC amplifying by PCR two double-stranded DNA fragments which introduce a  
CC linker sequence and a mutant sequence at each terminal region, using the  
CC gene encoding a protein as the template, connecting each double-stranded  
CC DNA fragment and amplifying a linear double-stranded DNA and thus  
CC producing a mutant protein. The polymorphisms present within an  
CC individual influence the sensitivity of that individual with respect to  
CC certain medical treatments since the variant protein will often exhibit  
CC varied function, with respect to the wild-type. Thus production and  
CC analysis of the structure and activity of such variants are often  
CC extremely useful. The method of the invention may be useful for producing  
CC a variant or mutant protein using a cell-free protein synthesis system.  
CC The mutant protein may be useful in determining the responsiveness to a  
CC medical agent, disease correlation, etc., as well as for analyzing  
CC protein function. The current sequence is that of the His tag peptide  
CC which was used in an exemplification of the invention.  
XX  
SQ Sequence 20 AA;

Query Match 100.0%; Score 95; DB 9; Length 20;  
Best Local Similarity 100.0%; Pred. No. 5.1e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 HLHNHVKKEHHAAHN 16  
Db 4 HLHNHVKKEHHAAHN 19  
RESULT 9  
AAV44209  
ID AAV44209 standard; peptide; 32 AA.  
XX  
AC AAV44209;  
XX  
DT 07-FEB-2000 (first entry)  
XX  
DE Fusion protein comprising affinity peptide and protein of interest.  
XX  
KM Fusion protein; secretory signal sequence; expression; vector;  
KM affinity peptide; nickel ion; protease cleavage site; host cell;  
KM recombinant fusion protein; yield; immobilization; purification;  
KM interaction; metal chelate resin; resin-fusion protein complex; elute;  
KM cleave; biological activity.  
XX  
OS Synthetic.  
XX  
PN WO957992-A1.  
XX  
PD 18-NOV-1999.  
XX  
PF 14-MAY-1999; 99WO-US010662.  
XX  
PR 14-MAY-1998; 98US-00078687.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PT Tthaga G, Jakhadze GG; ✓  
XX  
DR WPI; 2000-038923/03.  
XX  
PT Novel protein purification and immobilization methods based on a metal  
PT ion affinity site.  
XX  
PS Claim 6; Page 4; 42pp; English.  
XX  
CC The present sequence is a fusion protein. It comprises of a secretory  
CC signal sequence and an affinity peptide fused to the N-terminus of a  
CC protein of interest. A protease cleavage site is introduced between the  
CC affinity peptide, that has high affinity for nickel ions and the protein  
CC of interest. A vector containing the nucleic acid sequence encoding  
CC recombinant fusion protein can be introduced into host cells for its  
CC expression. Immobilization and purification of fusion proteins can be  
CC achieved via interaction of the fusion protein with metal chelate resin  
CC to form resin-fusion protein complex, which is then eluted and the  
CC affinity peptide cleaved from the protein of interest. This method is  
CC simple and provides high yield of purified proteins which retain full  
CC biological activity  
XX  
SQ Sequence 32 AA;  
Query Match 100.0%; Score 95; DB 3; Length 32;  
Best Local Similarity 100.0%; Pred. No. 8.9e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 HLHNHVKKEHHAAHN 16  
Db 5 HLHNHVKKEHHAAHN 20  
RESULT 10  
AAV44208

```

ID  AAY44208 standard; protein; 45 AA.
XX
AC  AAY44208;
XX
DT  07-FEB-2000 (first entry)
XX
DE  Protein comprising affinity peptide and multiple cloning site of vector
XX  pUC19/HS.
XX
KM  Vector pUC19/HS; fusion protein; multiple cloning site; translation site;
KM  secretory signal sequence; expression; affinity peptide; nickel ion;
KM  enterokinase cleavage site; host cell; recombinant fusion protein; yield;
KM  immobilization; purification; interaction; metal chelate resin;
KM  resin-fusion protein complex; elute; cleave; biological activity.
XX
OS  Synthetic.
XX
PN  WO957992-A1.
XX
PD  18-NOV-1999.
XX
PF  14-MAY-1999; 99WO-US010662.
XX
PR  14-MAY-1998; 98US-00078687.
XX
PA  (CLON-) CLONTECH LAB INC.
XX
PI  Tehaga G, Jokhade GQ;
XX
DR  WPI; 2000-038923/03.
XX
DR  N-PSDB; AA229088.
XX
PT  Novel protein purification and immobilization methods based on a metal
XX  ion affinity site.
XX
PS  Example 5; Page 37; 42pp; English.
XX
CC  The present sequence is a protein comprising the start of translation
CC  site, affinity peptide and the multiple cloning site of pUC19/HS vector
CC  encoding a fusion protein. The fusion protein comprises of a secretory
CC  signal sequence and an affinity peptide fused to the N-terminus of a
CC  protein of interest. An enterokinase cleavage site is introduced between
CC  the affinity peptide, that has high affinity for nickel ions and the
CC  protein of interest. This vector can be introduced into host cells for
CC  expression of recombinant fusion proteins. Immobilization and
CC  purification of fusion proteins can be achieved via interaction of the
CC  fusion protein with metal chelate resin to form resin-fusion protein
CC  complex, which is then eluted and the affinity peptide cleaved from the
CC  protein of interest. This method is simple and provides high yield of
CC  purified proteins which retain full biological activity
XX
SQ  Sequence 45 AA;

Query Match      100.0%; Score 95; DB 3; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 HLHNVHKEBHAAHN 16
    |||||
DB  11 HLHNVHKEBHAAHN 26

RESULT 11
AD267254      1
ID  AD267254 standard; protein; 211 AA.
XX
AC  AD267254;
XX
DT  30-JUN-2005 (first entry)
XX
DE  Escherichia coli S30 containing expression vector, PK7-CAT, protein.
XX  protein production; protein structure; analysis; X-ray crystallography.
XX

```

```

XX  Escherichia coli.
OS  Synthetic.
XX
PN  JP2005102513-A.
XX
PD  21-APR-2005.
XX
PF  26-SEP-2003; 2003JP-00336365.
XX
PR  26-SEP-2003; 2003JP-00336365.
XX
PA  (DOKU-) DOKURITSU GYOSAI HOJIN RIKAGAKU KENKYUSH.
XX  WPI; 2005-300044/31.
XX
DR  N-PSDB; AD267253.
XX
PT  Producing protein by cell free protein synthesis, involves using reaction
PT  liquid mixture containing salt of L-amino acid, D-amino acid and/or
XX  polyamino acid for carrying out isotopic labeling.
XX
PS  Disclosure; SEQ ID NO 4; 13pp; Japanese.
XX
CC  The invention relates to a novel method for producing a protein by cell
CC  free protein synthesis. The method involves using a reaction liquid
CC  mixture containing the salt of an L-amino acid, D-amino acid and/or
CC  polyamino acid for carrying out isotopic labeling, or using a reaction
CC  liquid mixture containing salt of a D-amino acid and/or polyamino acid.
CC  The invention further comprises: an isotope labeled protein produced by
CC  the above method; a reaction liquid mixture for cell free protein
CC  synthesis, comprising an L-amino acid that enables isotopic labeling to
CC  the S30 extract of Escherichia coli, buffer, salt and nucleotide
CC  triphosphoric acid, where the reaction liquid mixture contains 100-300 mM
CC  of potassium salt of D-amino acid as the salt; and a kit for cell-free
CC  protein synthesis, comprising the above mentioned reaction liquid
CC  mixture. The isotope labeled protein produced by the method is useful in
CC  carrying out three-dimensional structure analysis of the protein by NMR
CC  measurement. The isotope labeled protein produced by the method is useful
CC  in analyzing the enzyme activity, DNA binding ability and molecular
CC  weight of the protein, and in structural analysis of the protein by X-ray
CC  crystal structure analysis. This sequence represents an Escherichia coli
XX  S30 containing expression vector, PK7-CAT, protein of the invention.
XX

SQ  Sequence 211 AA;

Query Match      100.0%; Score 95; DB 9; Length 211;
Best Local Similarity 100.0%; Pred. No. 8.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 HLHNVHKEBHAAHN 16
    |||||
DB  4 HLHNVHKEBHAAHN 19

RESULT 12
AAY44207
ID  AAY44207 standard; protein; 278 AA.
XX
AC  AAY44207;
XX
DT  07-FEB-2000 (first entry)
XX
DE  Affinity peptide-GFPuv fusion protein.
XX
KM  Vector pGFPuv/HAT; fusion protein; secretory signal sequence; expression;
KM  affinity peptide; Green Fluorescent protein-UV mutant; GFPuv; nickel ion;
KM  protease cleavage site; host cell; recombinant fusion protein; yield;
KM  immobilization; purification; interaction; metal chelate resin;
KM  resin-fusion protein complex; elute; cleave; biological activity.
XX
OS  Synthetic.
XX
PN  WO957992-A1.
XX

```

XX 18-NOV-1999.  
 PD 14-MAY-1999; 99WO-US010662.  
 XX 14-MAY-1999; 98US-00078687.  
 PR 14-MAY-1998; 98US-00078687.  
 XX (CLON-) CLONTECH LAB INC.  
 XX Tchaga G, Jokhadze GG;  
 PI WPI; 2000-038923/03.  
 DR N-PSDB; AAZ29086.  
 XX Novel protein purification and immobilization methods based on a metal  
 PT ion affinity site.  
 XX Example 5; Page 32-33; 42pp; English.  
 XX The present sequence is a fusion protein encoded by partial cDNA of  
 CC vector pGPuv/HAT. The fusion protein comprises of a secretory signal  
 CC sequence and an affinity peptide fused to the N-terminus of Green  
 CC Fluorescent protein-UV mutant (GFPuv). A protease cleavage site is  
 CC introduced between the affinity peptide, that has high affinity for  
 CC nickel ions and the protein of interest. This vector can be introduced  
 CC into host cells for expression of recombinant fusion proteins.  
 CC Immobilization and purification of fusion proteins can be achieved via  
 CC interaction of the fusion protein with metal chelate resin to form resin-  
 CC fusion protein complex, which is then eluted and the affinity peptide  
 CC cleaved from the protein of interest. This method is simple and provides  
 CC high yield of purified proteins which retain full biological activity  
 XX Sequence 278 AA;  
 SQ  
 Query Match 100.0%; Score 95; DB 3; Length 278;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HLHNVKESHAHAN 16  
 Db 13 HLHNVKESHAHAN 28  
 RESULT 13  
 ABU08445  
 ID ABU08445 standard; protein; 340 AA.  
 XX  
 AC ABU08445;  
 DT 17-JUN-2003 (first entry)  
 DE Recombinant enterokinase (EK) fusion protein encoded by vector pHAT-EK.  
 XX  
 KM Metal ion affinity peptide; fusion protein; protein purification;  
 KM metal chelating resin; metal ion chelate resin; hard metal; Fe3+; Ca2+;  
 KM Al3+; Co2+; Ni2+; Zn2+; Co2+; protein-protein interaction;  
 KM metal ion affinity peptide-tagged recombinant protein; IMAC;  
 KM DNA-protein interaction; immobilised metal ion affinity chromatography;  
 KM gene expression; phosphorylation state; vector pHAT-EK; enterokinase.  
 XX  
 OS Unidentified.  
 OS Synthetic. 1  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 22  
 FT /note= "Encoded by TGA"  
 FT Misc-difference 330  
 FT /note= "Encoded by TGA"  
 XX  
 XX US2002164718-A1  
 XX 07-NOV-2002.

PF 15-MAY-2001; 2001US-00858332.  
 XX 25-SEP-1998; 98US-0101667P.  
 PR 23-SEP-1999; 98US-00404017.  
 XX (TCHG/) TCHAGA G S.  
 PA (JOKH/) JOKHADZE G G. ✓  
 XX Tchaga GS, Jokhadze GG;  
 PI WPI; 2003-361747/34.  
 DR N-PSDB; ABX94278.  
 XX New metal ion affinity peptide useful, when fused to a fusion partner  
 PT polypeptide, for protein purification methods and to study protein-  
 PT protein interactions and nucleic acid-protein interactions.  
 XX Example 1; Fig 2; 23pp; English.  
 XX The present invention relates to metal ion affinity peptides, fusion  
 CC proteins containing metal ion affinity peptides, and polynucleotide  
 CC sequences encoding the fusion proteins. The presence of a metal ion  
 CC affinity peptide in a fusion protein allows purification of the fusion  
 CC protein on a metal chelating resin. The method involves contacting a  
 CC sample comprising a fusion protein with a metal ion chelate resin  
 CC comprising a first metal ion, preferably a hard metal ion such as Fe3+,  
 CC Ca2+ or Al3+ and eluting any resultant bound fusion protein from the  
 CC resin. The resin comprises an immobilised Co2+ ion. The method further  
 CC comprises contacting the sample with a second immobilised metal ion  
 CC affinity resin comprising a second immobilised metal ion and eluting any  
 CC resultant bound fusion protein from the first and second resins. The  
 CC second metal ion is an intermediate metal ion such as Cu2+, Ni2+, Zn2+ or  
 CC Co2+. The metal ion affinity peptide-tagged recombinant proteins are  
 CC useful for the study of protein-protein interactions and nucleic acid  
 CC molecule-protein interactions, using solid phase immobilised metal ion  
 CC affinity chromatography (IMAC). They are also useful in high throughput  
 CC systems which find use in massive parallel gene expression experiments,  
 CC e.g. to determine the effect of an agent on synthesis of a protein or set  
 CC of proteins, to analyse developmental stage-specific, or tissue-specific  
 CC synthesis of a protein and to analyse the phosphorylation state of a  
 CC protein. These methods find use in applications to characterise a protein  
 CC of unknown identity or function, and in enzymatic reactions. The present  
 CC sequence represents the recombinant enterokinase (EK) fusion protein  
 CC encoded by vector pHAT-EK  
 XX  
 SQ Sequence 340 AA;  
 Query Match 100.0%; Score 95; DB 6; Length 340;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HLHNVKESHAHAN 16  
 Db 71 HLHNVKESHAHAN 86  
 RESULT 14  
 ADX26604  
 ID ADX26604 standard; peptide; 19 AA.  
 XX  
 AC ADX26604;  
 DT 21-APR-2005 (first entry)  
 DE Histidine tag, HAT, peptide, SEQ ID 14.  
 XX Fusion protein; protein engineering.  
 XX Synthetic.  
 OS  
 OS US2005032173-A1.  
 XX 10-FEB-2005.

```
XX 05-AUG-2003; 2003US-00634645.
PF 05-AUG-2003; 2003US-00634645.
PR 05-AUG-2003; 2003US-00634645.
XX (ROJA/) ROJAS M.
PA (MORA/) MORA A D.
XX
XX Rojae M, Mora AL;
PI
XX MPI; 2005-141725/15.
DR
XX
XX Novel fusion protein having membrane-translocating sequence and Ikappab
PT protein, useful for treating immune-related diseases such as allergy, and
PT apoptosis-related disorder such as cancer.
XX
XX Claim 11; SEQ ID NO 14; 31pp; English.
XX
XX The present invention relates to novel fusion proteins (I) comprising a
CC membrane-translocating sequence (MTS, ADX26591-ADX26599) and an Ikappab
CC protein. (I) further comprises a tag amino acid sequence or protein
CC (ADX26600-ADX26608) for attaching the MTS to the Ikappab protein. (I) is
CC useful for treating an immune-related disorder, an apoptosis-related
CC disorder such as cancer, or preventing an immune response in a host,
CC which involves administering (I). The immune response is associated with
CC at least one type of an allergy, asthma, contact dermatitis, delayed-type
CC hypersensitivity, a wound-healing, allergic rhinitis, food
CC hypersensitivity, ectopic dermatitis, inflammatory bowel disease, an
CC immunological disease of the lung, eosinophilic pneumonia, idiopathic
CC pulmonary fibrosis, hypersensitivity pneumonitis, an autoimmune or immune
CC -mediated skin disease, bullous skin disease, erythema multiforme,
CC psoriasis, gluten-sensitive enteropathy, Whipple's disease, systemic
CC lupus erythematosus, rheumatoid arthritis, osteoarthritis, juvenile
CC chronic arthritis, ankylosing spondylitis, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's disease, pleuritis,
CC sarcoidosis, amyloidosis, autoimmune hemolytic anemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, myasthenia gravis, a demyelinating disease of the central or
CC peripheral nervous system, idiopathic demyelinating polyneuropathy,
CC Guillain-Barre syndrome, a chronic inflammatory demyelinating
CC polyneuropathy, a hepatobiliary disease, an infectious or autoimmune
CC chronic active hepatitis, primary biliary cirrhosis, granulomatous
CC hepatitis, sclerosing cholangitis, Graves disease, a transplantation-
CC associated disease, a graft rejection, and graft-versus-host disease.
XX
XX Sequence 19 AA;
SQ
XX
XX Query Match 91.6%; Score 87; DB 9; Length 19;
XX Best Local Similarity 93.8%; Pred. No. 8.6e-07;
XX Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 HLHNVKKEHAAHN 16
QY |||||
DB 3 HLHNVKKEHAAHN 18
XX
XX RESULT 15
XX ADG14415
ID ADG14415 standard; peptide; 19 AA.
XX
XX ADG14415;
AC
XX
XX 26-FEB-2004 (first entry)
DT
XX
XX Chicken lactate dehydrogenase peptide.
DE
XX
XX Chicken; lactate dehydrogenase ; Green Fluorescent Protein; GFP.
XX
XX Gallus gallus.
OS
XX WO2003072143-A1.
PN
XX 04-SEP-2003.
PD
```

```
XX 27-FEB-2003; 2003WO-US005937.
PF 27-FEB-2003; 2003WO-US005937.
PR 27-FEB-2002; 2002US-0360350P.
XX (PHAR-) PHARMAIN LTD.
PA
XX
XX Bolotin EM;
PI
XX MPI; 2003-803793/75.
DR
XX
XX Biocompatible composition useful for the treatment of disease e.g.
PT bacterial infection comprises a carrier with a metal binding domain,
PT metal ion chelated to the metal binding domain of the carrier and an
PT active agent.
XX
XX Example 7; Page 63; 77pp; English.
XX
XX The present invention relates to a biocompatible composition (CI)
CC comprises a carrier with metal binding domain (MBD), a metal ion chelated
CC to MBD of the carrier, and an active agent (CI) with MBD chelated to the
CC metal ion. CI is useful in the manufacture of a medicament for the
CC treatment of diseases or condition e.g. infection caused by bacteria,
CC viruses (e.g. HIV, herpes, hepatitis) and pathogenic fungi (Candida sp
CC ). The present sequence is a peptide fragment of chicken lactate
CC dehydrogenase used in the construction of a His-tagged Green Fluorescent
CC protein (GFP) variant to illustrate binding to nitrilotriacetic acid(NTA)
CC -linked co-polymers in the presence of metal ions.
XX
XX Sequence 19 AA;
SQ
XX
XX Query Match 81.1%; Score 77; DB 7; Length 19;
XX Best Local Similarity 80.0%; Pred. No. 3.2e-05;
XX Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 HLHNVKKEHAAHN 15
QY |||||
DB 3 HLHNVKKEHAAHN 17
XX
XX RESULT 16
XX ABP03327
ID ABP03327 standard; protein; 69 AA.
XX
XX ABP03327;
AC
XX
XX 24-JUN-2002 (first entry)
DT
XX
XX Human ORF protein sequence SEQ ID NO:6636.
DE
XX
XX Human; open reading frame; ORF; gene therapy; cancer; cirrhosis;
XX hyperproliferative disorder; psoriasis; benign tumor; haemorrhage;
XX degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX hypertension; hypothyroidism; cholesterol ester storage disease;
XX immune deficiency; immune disorder; infectious disease;
XX autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
XX myasthenia gravis.
XX
XX Homo sapiens.
OS
XX
XX MO200192523-A2.
PN
XX
XX 06-DEC-2001.
PD
XX
XX 29-MAY-2001; 2001WO-US010836.
PF
XX
XX 30-MAY-2000; 2000US-0206132P.
PR
XX 29-AUG-2000; 2000US-0228716P.
XX
XX (CURA-) CURAGEN CORP.
PA
XX
XX Shinketsu RA, Leach MD;
PI
```

XX WPI; 2002-106308/14.  
DR N-PSDB; ABLN19079.  
XX  
XX Novel human polypeptides and polynucleotides useful for diagnosing,  
PT preventing and treating cardiovascular disease, neurodegenerative,  
PT hyperproliferative disorders and autoimmune disorders.  
XX  
PS Disclosure; SEQ ID NO 6636; 1037bp; English.  
XX  
XX The present invention describes substantially purified human proteins  
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
CC in the specification). ABLN15762 to ABLN27252 encode the human ORFX  
CC proteins given in ABLN00010 to ABLN1500. ORFX proteins are useful for  
CC treating or preventing a pathology associated with an ORFX-associated  
CC disorder in humans, and in the manufacture of a medicament for treating a  
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
CC sequences can be used in gene therapy. ORFX sequences can be used in the  
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
CC psoriasis, benign tumors, keloid, degenerative disorders, haemorrhage,  
CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
CC storage disease, various immune deficiencies and disorders, infectious  
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
CC bone degenerative disorders, or periodontal disease, and for gut  
CC protection or regeneration and treatment of lung or liver fibrosis,  
CC reperfusion injury in various tissues and conditions resulting from  
CC systemic cytokine damage. N.B. The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 69 AA;  
XX  
Query Match 60.0%; Score 57; DB 5; Length 69;  
Best Local Similarity 56.2%; Pred. No. 0.2;  
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 HLHNHVKHEHAH 16  
Db 31 HSIHYTHSTHATHN 46  
XX  
RESULT 17  
ABBB0446  
ID ABBB0446 standard; protein; 837 AA.  
XX  
AC ABBB0446;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 8130.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KM pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
XX  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW,

XX WPI; 2001-656860/75.  
DR N-PSDB; ABL04549.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX  
PS Disclosure; SEQ ID NO 8130; 21bp + Sequence listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 837 AA;  
XX  
Query Match 55.8%; Score 53; DB 4; Length 837;  
Best Local Similarity 53.3%; Pred. No. 16;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
QY 1 HLHNHVKHEHAH 15  
Db 318 HHNHAAQAQHAH 332  
XX  
RESULT 18  
ABBB4529  
ID ABBB4529 standard; protein; 402 AA.  
XX  
AC ABBB4529;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 20379.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KM pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
XX  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW,  
XX  
DR WPI; 2001-656860/75.  
XX  
DR N-PSDB; ABL08652.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX  
PS Disclosure; SEQ ID NO 20379; 21bp + Sequence listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of

CC Insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA  
CC sequences (ABU1840-ABU16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 402 AA;  
Query Match 53.7%; Score 51; DB 4; Length 402;  
Best Local Similarity 66.7%; Pred. No. 14;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 4 HNYHKEHHAAH 15  
DB 239 HHVHAHAHAHAH 250  
RESULT 19  
AAU1504  
ID AAG11504 standard; protein; 299 AA.  
XX  
AC AAG11504;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 10245.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridization assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135625P.  
PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.

PR	04-AUG-1999	99US-0147204P
PR	04-AUG-1999	99US-0147302P
PR	05-AUG-1999	99US-0147192P
PR	05-AUG-1999	99US-0147260P
PR	06-AUG-1999	99US-0147303P
PR	06-AUG-1999	99US-0147416P
PR	09-AUG-1999	99US-0147493P
PR	09-AUG-1999	99US-0147935P
PR	10-AUG-1999	99US-0148171P
PR	11-AUG-1999	99US-0148319P
PR	12-AUG-1999	99US-0148431P
PR	13-AUG-1999	99US-0148565P
PR	13-AUG-1999	99US-0148664P
PR	16-AUG-1999	99US-0149368P
PR	17-AUG-1999	99US-0149175P
PR	23-AUG-1999	99US-0149930P
PR	25-AUG-1999	99US-0150566P
PR	26-AUG-1999	99US-0150884P
PR	27-AUG-1999	99US-0151065P
PR	27-AUG-1999	99US-0151066P
PR	27-AUG-1999	99US-0151083P
PR	30-AUG-1999	99US-0151103P
PR	31-AUG-1999	99US-0151438P
PR	01-SEP-1999	99US-0151930P
PR	07-SEP-1999	99US-0152330P
PR	10-SEP-1999	99US-0152367P
PR	13-SEP-1999	99US-0153070P
PR	15-SEP-1999	99US-0153578P
PR	15-SEP-1999	99US-0154018P
PR	16-SEP-1999	99US-0154039P
PR	20-SEP-1999	99US-0154779P
PR	22-SEP-1999	99US-0155139P
PR	23-SEP-1999	99US-0155486P
PR	24-SEP-1999	99US-0155659P
PR	28-SEP-1999	99US-0156458P
PR	29-SEP-1999	99US-0156569P
PR	04-OCT-1999	99US-0157117P
PR	05-OCT-1999	99US-0157753P
PR	06-OCT-1999	99US-0157865P
PR	07-OCT-1999	99US-0158029P
PR	08-OCT-1999	99US-0158232P
PR	12-OCT-1999	99US-0158393P
PR	13-OCT-1999	99US-0159293P
PR	13-OCT-1999	99US-0159294P
PR	13-OCT-1999	99US-0159295P
PR	14-OCT-1999	99US-0159329P
PR	14-OCT-1999	99US-0159330P
PR	14-OCT-1999	99US-0159331P
PR	14-OCT-1999	99US-0159637P
PR	18-OCT-1999	99US-0159684P
PR	18-OCT-1999	99US-0159984P
PR	21-OCT-1999	99US-0160741P
PR	21-OCT-1999	99US-0160767P
PR	21-OCT-1999	99US-0160778P
PR	21-OCT-1999	99US-0160780P
PR	21-OCT-1999	99US-0160814P
PR	21-OCT-1999	99US-0160815P
PR	22-OCT-1999	99US-0160980P
PR	22-OCT-1999	99US-0160981P
PR	22-OCT-1999	99US-0160989P
PR	25-OCT-1999	99US-0161404P
PR	25-OCT-1999	99US-0161405P
PR	25-OCT-1999	99US-0161406P
PR	26-OCT-1999	99US-0161559P
PR	26-OCT-1999	99US-0161560P
PR	26-OCT-1999	99US-0161616P
PR	28-OCT-1999	99US-0161320P
PR	28-OCT-1999	99US-0161992P
PR	28-OCT-1999	99US-0161993P

Query Match	Best Local Similarity	Score 50;	DB 3;	Length 299;
Matches	8;	Conservative	2;	Mismatches
			Indels	Gaps
QY	1	HLINVKEHEHNAHN 16	6;	0;
DB	178	HDHNEHEHEHEHH 193		
RESULT 20				
AGG11503				
ID	AGG11503	standard; protein; 300 AA.		
XX	AC	AGG11503;		
XX	DT	17-OCT-2000 (first entry)		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 10244.			
XX	Protein identification; signal transduction pathway; metabolic pathway;			
KW	hybridisation assay; genetic mapping; gene expression control; promoter;			
KX	termination sequence.			
XX	Arabidopsis thaliana.			
XX	OS	Arabidopsis thaliana.		
XX	PN	EP1033405-A2.		
XX	PD	06-SEP-2000.		
XX	PF	25-FEB-2000; 2000EP-00301439.		
XX	PR	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999;	99US-0123180P.		
PR	09-MAR-1999;	99US-0123548P.		
PR	23-MAR-1999;	99US-0125788P.		
PR	25-MAR-1999;	99US-0126264P.		
PR	29-MAR-1999;	99US-0126785P.		
PR	01-APR-1999;	99US-0127452P.		
PR	06-APR-1999;	99US-0128234P.		
PR	08-APR-1999;	99US-0128714P.		
PR	16-APR-1999;	99US-0129845P.		
PR	19-APR-1999;	99US-0130077P.		
PR	23-APR-1999;	99US-0130449P.		
PR	23-APR-1999;	99US-0130510P.		
PR	23-APR-1999;	99US-0130891P.		
PR	28-APR-1999;	99US-0131449P.		
PR	30-APR-1999;	99US-0132048P.		
PR	30-APR-1999;	99US-0132407P.		
PR	04-MAY-1999;	99US-0132484P.		
PR	05-MAY-1999;	99US-0132485P.		
PR	06-MAY-1999;	99US-0132486P.		
PR	07-MAY-1999;	99US-0132487P.		
PR	11-MAY-1999;	99US-0132663P.		
PR	14-MAY-1999;	99US-0134256P.		
PR	14-MAY-1999;	99US-0134218P.		
PR	14-MAY-1999;	99US-0134219P.		
PR	14-MAY-1999;	99US-0134221P.		
PR	18-MAY-1999;	99US-0134370P.		
PR	19-MAY-1999;	99US-0134768P.		
PR	20-MAY-1999;	99US-0134941P.		
PR	21-MAY-1999;	99US-0135124P.		
PR	24-MAY-1999;	99US-0135353P.		
PR	25-MAY-1999;	99US-0135629P.		
PR	27-MAY-1999;	99US-0136021P.		
PR	28-MAY-1999;	99US-0136392P.		
PR	01-JUN-1999;	99US-0136788P.		
PR	03-JUN-1999;	99US-0137222P.		
PR	04-JUN-1999;	99US-0137528P.		
PR	07-JUN-1999;	99US-0137502P.		
PR	08-JUN-1999;	99US-0137724P.		
PR	09-JUN-1999;	99US-0138094P.		



PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 22-JUN-1999; 99US-0139817P.  
PR 23-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149375P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 23-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149920P.  
PR 25-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-015138P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-015717P.  
PR 05-OCT-1999; 99US-015753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 52.6%; Score 50; DB 3; Length 300;  
Best Local Similarity 50.0%; Pred. No. 14;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
QY 1 HIIHVHKEEHAHIN 16

Db 179 HDHNEHEHEHEHH 194

## RESULT 21

AA24393  
ID AAR24393 standard; protein; 351 AA.

XX AAR24393;

XX 25-MAR-2003 (revised)

DT 22-NOV-1992 (first entry)

XX Sequence of Histidine-rich protein (HicRp).

XX Malaria vaccine; Histidine-rich protein; cytoadherence.

XX Plasmodium lophurae.

XX Key Location/Qualifiers

FT Peptide 1..23 /label= signal

FT Peptide 24..47 /label= pro-peptide

FT Modified-site 40..42 /label= potential glycosylation site

XX US5116965-A.

XX 26-MAY-1992.

XX 26-AUG-1986; 86US-00900401.

XX 26-AUG-1986; 86US-00900401.

XX (SLOK ) SLOAN KETTERING INST CANCER.

XX Ravetch JV, Pologe L;

XX WPI; 1992-199590/24.

XX N-PSDB; AAQ24393.

XX Histidine-rich protein associated with plasmodium knob phenotype, and DNA encoding it - useful for in vitro diagnosis of P. falciparum infection.

XX Disclosure; Fig 7A-B; 29pp; English.

XX Two variants of HicRp are produced by P. falciparum. One is associated with what is referred to as "knobby phenotype" (K30) and "knobless phenotype" (K-). The "knobby" and "knobless" phenotypes have been implicated in cytoadherence, which is characteristic of erythrocyte infection. It has now been found that cDNA expressing both K+ and K- HicRp can be obtained by the use of P. lophurae HicRp expressing DNA. The genomic clone (AAQ25532) is encoded in two exons, separating the signal peptide-encoding sequence from the pro-sequence, confirming that synthesis of the protein occurs via the preproprotein. Oligo. probes synthesised to the signal peptide-encoding exon reveal multiple homologous DNA sequences in the P. lophurae genome. The sequence of mature proteins is arranged in numerous tandem repeats with up to nine histidine residues in a row, similar to other Plasmodium proteins for which sequence data have so far been reported. (Updated on 25-MAR-2003 to correct PD field.)

XX Sequence 351 AA;

Query Match 52.6%; Score 50; DB 2; Length 351;

Best Local Similarity 50.0%; Pred. No. 17;

Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 HLILNVKKEKHAHVN 16

Db 123 HAAHHNHNEHEHHN 138

## RESULT 22

AA24393  
ID AAG11502 standard; protein; 448 AA.

XX AAG11502;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 10243.

XX Protein identification: signal transduction pathway; metabolic pathway;

XX hybridisation assay; genetic mapping; gene expression control; promoter;

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 08-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130449P.

XX 23-APR-1999; 99US-0130510P.

XX 28-APR-1999; 99US-0130891P.

XX 30-APR-1999; 99US-0131499P.

XX 04-MAY-1999; 99US-0132407P.

XX 05-MAY-1999; 99US-0132484P.

XX 06-MAY-1999; 99US-0132486P.

XX 06-MAY-1999; 99US-0132487P.

XX 07-MAY-1999; 99US-0132863P.

XX 11-MAY-1999; 99US-0134218P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134221P.

XX 18-MAY-1999; 99US-0134370P.

XX 19-MAY-1999; 99US-0134768P.

XX 20-MAY-1999; 99US-0134941P.

XX 21-MAY-1999; 99US-0135124P.

XX 24-MAY-1999; 99US-0135353P.

XX 25-MAY-1999; 99US-0135629P.

XX 27-MAY-1999; 99US-0136021P.

XX 28-MAY-1999; 99US-0136392P.

XX 01-JUN-1999; 99US-0137222P.

XX 03-JUN-1999; 99US-0137528P.

XX 04-JUN-1999; 99US-0137502P.

XX 07-JUN-1999; 99US-0137724P.

XX 08-JUN-1999; 99US-0138094P.

XX 10-JUN-1999; 99US-0138540P.

XX 14-JUN-1999; 99US-0138847P.

XX 16-JUN-1999; 99US-0139119P.

XX 17-JUN-1999; 99US-0139452P.

XX 18-JUN-1999; 99US-0139453P.

XX 18-JUN-1999; 99US-0139452P.

XX 18-JUN-1999; 99US-0139454P.

XX 18-JUN-1999; 99US-0139455P.

XX 18-JUN-1999; 99US-0139456P.

XX 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 12-JUL-1999; 99US-0142220P.  
PR 13-JUL-1999; 99US-0142977P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144255P.  
PR 19-JUL-1999; 99US-0144311P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 22-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.

PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151338P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153707P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 52.6%; Score 50; DB 3; Length 448;  
Best Local Similarity 50.0%; Pred. No. 23;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 HLIHNVKKEEHAHNM 16  
Db 327 HDHNNHEHEHEHEHH 342

RESULT 23  
ADW13460  
ID ADW13460 standard; protein; 198 AA.  
XX  
AC ADW13460;  
XX

DT 07-APR-2005 (first entry)  
 XX E. tenella oocyst sporocyst protein EtOS22.  
 DE antiparasitic; Vaccine; oocyst sporocyst; coccidiosis; EtOS22.  
 XX Eimeria tenella.  
 OS WO2005005472-A1.  
 PN 20-JAN-2005.  
 PD 30-JUN-2004; 2004MO-EP007080.  
 PF 04-JUL-2003; 2003DE-01030235.  
 PR (FARB ) BAYER HEALTHCARE AG.  
 PA Grefl G, Hose R, Kruecken J, Wunderlich F;  
 PI WPI; 2005-092065/10.  
 DR N-PSDB; ADW13459, ADW13461.  
 DR New Eimeria polynucleotides and corresponding EtOS22 proteins useful for  
 PT diagnosing, preventing or treating infections caused by Eimeria (e.g.  
 PT coccidiosis) or for finding drugs that may treat or prevent coccidiosis.  
 XX Claim 1; SEQ ID NO 2; 79pp; English.  
 XX The invention relates to a polynucleotide encoding an oocyst sporocyst  
 CC protein belonging to the parasite of the species Eimeria tenella. The  
 CC composition (including the polynucleotide, polypeptide, vector or  
 CC expression system, or antibody) is useful for producing a vaccine. The  
 CC active compound that modulates the activity of the EtOS22 protein is used  
 CC for producing a drug for treating coccidiosis, or the antibody is used as  
 CC parasiticides. The composition and methods may also be used for  
 CC diagnosing, preventing or treating coccidiosis or for finding drugs that  
 CC may treat or prevent coccidiosis. The present sequence represents the  
 CC amino acid sequence of E. tenella oocyst sporocyst protein EtOS22.  
 XX  
 XX Sequence 198 AA;  
 SQ  
 Query Match 52.1%; Score 49.5; DB 9; Length 198;  
 Best Local Similarity 58.8%; Pred. No. 10;  
 Matches 10; Conservative 1; Mismatches 5; Indels 1; Gaps 1;  
 QY 1 HLHNHVKKEHANA-HN 16  
 Db 145 HHEHNHVPQHQAQHN 161  
 RESULT 24  
 ADT87823  
 ID ADT87823 standard; protein; 334 AA.  
 AC ADT87823;  
 DT 13-JAN-2005 (first entry)  
 XX Plant homologue of yeast SRP YII023C #2.  
 DE Plant homologue of yeast SRP YII023C #2.  
 XX  
 KW Plant; Stress-related protein; SRP; environmental stress; abiotic stress;  
 KW drought; heat; cold; salt.  
 OS Oryza sativa.  
 OS WO2004093398-A2.  
 PN 28-OCT-2004.  
 PD 15-APR-2004; 2004MO-US011888.  
 PF 15-APR-2003; 2003EP-0008080.  
 PR

PR 02-MAY-2003; 2003EP-00039728.  
 PR 01-AUG-2003; 2003EP-00016672.  
 PR 30-SEP-2003; 2003EP-00022225.  
 XX  
 XX (BADI ) BASF PLANT SCI GMBH.  
 PA Puzio P, Chardonnens A, Shirley A, Wang X, Sarria-Willan R;  
 PI Mckersie B, Chen R;  
 DR WPI; 2004-766863/75.  
 DR N-PSDB; ADT87821.  
 XX  
 XX New isolated nucleic acid molecule comprises a sequence encoding Stress-  
 PT Related Protein (SRP), useful for producing transformed plants with  
 PT altered metabolic activity resulting in increased tolerance or resistance  
 PT to environmental stress.  
 XX Claim 20; Page 899; 91pp; English.  
 XX The invention relates an isolated nucleic acid molecule comprises a  
 CC nucleic acid molecule encoding a Stress-Related Protein (SRP), from Yeast  
 CC or E. coli and their homologues from Rice, Soybean and Rape. Also  
 CC included are a transformed plant cell with altered metabolic activity  
 CC compared to a corresponding non-transformed wild type plant cell (where  
 CC the metabolic activity is altered by transformation with a SRP coding  
 CC nucleic acid and results in increased tolerance and/or resistance to an  
 CC environmental stress as compared to a corresponding non-transformed wild  
 CC type plant cell), a transgenic plant generated from the plant cell above  
 CC (and which is a monocot or dicot plant, or a gymnosperm plant), a seed  
 CC produced by a transgenic plant above (where the seed is genetically  
 CC homozygous for a transgene conferring altered metabolic activity  
 CC resulting in an increased tolerance to environmental stress as compared  
 CC to a corresponding non-transformed wild type plant), a nucleic acid  
 CC construct which confers the expression of the nucleic acid molecule above  
 CC (comprising one or more regulatory elements, where expression of the SRP  
 CC coding nucleic acid in a host cell results in altered metabolic activity  
 CC resulting in increased tolerance to environmental stress as compared to a  
 CC corresponding non-transformed wild type host cell), a vector comprising  
 CC the nucleic acid molecule above or the nucleic acid construct, a host  
 CC cell which has been transformed stably or transiently with the vector (or  
 CC the nucleic acid molecules above, or the nucleic acid construct), an  
 CC isolated stress Related Protein (SRP) selected from the amino acid  
 CC sequences fully given in the specification and/or its homologues, a  
 CC method of producing a transgenic plant with altered metabolic activity  
 CC compared to a corresponding non-transformed wild type plant cell,  
 CC modifying stress tolerance of a plant, detecting environmental stress in  
 CC plant cells or plants, screening plant cells or plants for increased  
 CC tolerance and/or resistance to environmental stress, breeding plant cells  
 CC or plants towards increased tolerance and/or resistance to environmental  
 CC stress, increasing tolerance of a plant to at least one abiotic stress,  
 CC a plant transformed with the nucleic acids above and a seed of the plant.  
 CC The altered metabolic activity and/or a SRP encoding nucleic acids or its  
 CC homologues are useful as markers for selection of plants or plant cells  
 CC with increased tolerance to environmental stress, or for detection of  
 CC stress in plants or plant cells. The nucleic acids are useful for  
 CC producing transformed plants with altered metabolic activity resulting in  
 CC increased tolerance and/or resistance to an environmental stress  
 CC (drought, heat, cold and salt) as compared to a corresponding non-  
 CC transformed wild-type plant cell. The present sequence is a plant SRP of  
 CC the invention.  
 XX  
 XX Sequence 334 AA;  
 SQ  
 Query Match 52.1%; Score 49.5; DB 8; Length 334;  
 Best Local Similarity 56.2%; Pred. No. 19;  
 Matches 9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;  
 QY 1 HLHNHVKKEHANA-HN 16  
 Db 36 HQNHN-HSHEHSHANS 50  
 RESULT 25

```

AAO10111
ID AAO10111 standard; protein; 117 AA.
XX
AC AAO10111;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 24003.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN MO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001MO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSR-) HYSRQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR N-PSDB; AAI90042.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 20; SEQ ID NO 24003; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 117 AA;

Query Match          51.6%; Score 49; DB 4; Length 117;
Best Local Similarity 46.7%; Pred. No. 6.6;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 HLIHVHKEEHAH 15
   ||:|:|:|:|:|
DB 69 HLVGMHKKKTH 83

RESULT 26
AAB33000
ID AAB33000 standard; protein; 134 AA.
XX
AC AAB33000;
XX
DT 25-JAN-2001 (first entry)
XX
DE Pinus radiata transcription factor protein sequence #127.
XX
KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;

```

```

KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW homeodomain zipper; LIM domain; AP2; ERBs; zinc finger domain;
KW type 2 Cys2His2; CCAAT box element; MYB.
XX
OS Pinus radiata.
XX
PN MO200053724-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000MO-US006112.
XX
PR 11-MAR-1999; 99US-00266513.
PR 18-AUG-1999; 99US-0149485P.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLEET-) FLETCHEE CHALLENGE FORESTS LTD.
XX
PI Wood M, McGrath A, Shenk MA, Glenn M;
XX
DR WPI; 2000-579369/54.
XX
PT New isolated polynucleotide encoding a plant transcription factor for
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT having modified gene expression or modified activity of a polypeptide.
XX
PS Claim 8; Page 386; 747pp; English.
XX
CC The present invention relates to novel plant transcription factors from
CC Eucalyptus grandis or Pinus radiata. The present sequence is one such
CC transcription factor. The transcription factor may be used to produce a
CC plant having modified gene expression such as a woody plant e.g. a
CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
CC to modify the activity of a polypeptide in a plant. The transcription
CC factors of the present invention are members from the following families
CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
CC zipper, LIM domain, AP2 and ERBs, zinc finger domains of type 2
CC Cys2His2, CCAAT box elements and MYB
XX
SQ Sequence 134 AA;

Query Match          51.6%; Score 49; DB 3; Length 134;
Best Local Similarity 37.5%; Pred. No. 7.8;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 HLIHVHKEEHAH 16
   ||:|:|:|:|:|
DB 96 HLLKNHRRKPVSHS 111

RESULT 27
ABB69867
ID ABB69867 standard; protein; 366 AA.
XX
AC ABB69867;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 36393.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN MO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001MO-US009231.
XX

```

PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
XX (PEKE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li FWD, Myers EW,  
XX  
XX WPI; 2001-656860/75.  
DR N-PSDB; ABL13970.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
PS  
XX Disclosure; SEQ ID NO 36393; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signaling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
XX ABB72072). The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at [http://www.wipo.int/pub/published\\_pct\\_sequences](http://www.wipo.int/pub/published_pct_sequences)  
XX  
SQ Sequence 368 AA;  
Query Match 51.6%; Score 49; DB 4; Length 368;  
Best Local Similarity 42.9%; Pred. No. 26;  
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
QY 3 HNNVKEEHAHNN 16  
:|:|:|:|:|:|  
Db 258 VHHSTHSHSHSHN 271  
RESULT 28  
ADT59295  
ID ADT59295 standard; protein; 392 AA.  
XX  
AC ADT59295;  
XX  
DT 13-JUN-2005 (first entry)  
XX  
DE Plant polypeptide, SEQ ID 9372.  
XX  
KM Plant; transgenic; cold tolerance; growth rate; drought tolerance;  
KM disease resistance; galactomannan production; plant growth regulator;  
KM heat tolerance; herbicide tolerance; lignin production;  
KM extreme osmotic tolerance; pathogen resistance;  
KM pest resistance; yield improvement; seed oil yield; seed protein yield.  
XX  
OS Viridiplantae.  
XX  
XX US2004216190-A1.  
XX  
XX 28-OCT-2004.  
XX  
PF 18-DEC-2003; 2003US-00739930.  
XX  
XX 28-APR-2003; 2003US-00424599.  
PR 28-APR-2003; 2003US-00425115.  
XX  
XX (KOVA/) KOVALIC D K.  
XX  
XX KOVALIC DK;  
XX  
XX WPI; 2004-757369/74.  
XX  
XX New recombinant DNA constructs useful in the field of biochemistry and  
XX genetics, and in particular for producing transgenic plants with improved

PT biological characteristics.  
XX  
XX Claim 2; SEQ ID NO 9372; 14pp; English.  
PS  
XX  
XX The invention relates a recombinant DNA construct comprising a  
XX polynucleotide having any of 5544 nucleotide sequences (CDNAs SEQ ID NO:  
XX 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences  
XX (SEQ ID NO: 5545-11088). The CDNAs and proteins are from corn, soybean,  
XX Arabidopsis, wheat and rape but the specification does not indicate which  
XX sequences is derived from which organism. Also included is a method of  
XX producing a plant having an improved property, comprising transforming a  
XX plant with a recombinant DNA construct comprising a promoter region  
XX functional in a plant cell operably joined to a polynucleotide encoding a  
XX polypeptide associated with the property, and growing the transformed  
XX plant. The property is selected from improving plant cold tolerance, for  
XX manipulating growth rate in plant cells by modification of the cell cycle  
XX pathway, for improving plant drought tolerance, for providing increased  
XX resistance to plant disease, for galactomannan production, for production  
XX of plant growth regulators, for improving plant heat tolerance, for  
XX improving plant tolerance to herbicides, for increasing the rate of  
XX homologous recombination in plants, for lignin production, for improving  
XX plant tolerance to extreme osmotic conditions, for improving plant  
XX tolerance to pathogens or pests, for yield improvement by modification of  
XX photosynthesis, for modifying seed oil yield and/or content, for  
XX modifying seed protein yield and/or content, for yield improvement by  
XX modification of carbohydrate, nitrogen or phosphorus use and/or uptake  
XX and for yield improvement by providing improved plant growth and  
XX development under at least one stress condition. The polynucleotide may  
XX also encode a plant transcription factor. The methods and compositions of  
XX the present invention are useful in the field of biochemistry and  
XX genetics, in particular for producing transgenic plants with improved  
XX biological characteristics such as increased heat, improved nitrogen  
XX flow, increasing plant tolerance to cold or heat, improving plant  
XX tolerance to extreme osmotic and drought conditions, and improving plant  
XX tolerance to plant pests or pathogens. They can also be used in physical  
XX arrays of molecules, plant breeding markers, computer-based storage and  
XX analysis systems. The present sequence is one of the 5544 plant protein  
XX sequences of the invention. Note: The sequence data for this patent did  
XX not form part of the printed specification, but was obtained in  
XX electronic format directly from USPTO at  
XX [seqdata.uspto.gov/sequence.html?docID=20040216190](http://seqdata.uspto.gov/sequence.html?docID=20040216190).  
XX  
SQ Sequence 392 AA;  
Query Match 51.6%; Score 49; DB 8; Length 392;  
Best Local Similarity 37.5%; Pred. No. 28;  
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;  
QY 1 HLNHVKKEHAHNN 16  
||:|:|:|:|:|  
Db 96 HLNKNIHRRKPVHSHS 111  
RESULT 29  
ADX68442  
ID ADX68442 standard; protein; 418 AA.  
XX  
XX ADX68442;  
XX  
XX 21-APR-2005 (first entry)  
XX  
DE Plant full length insert polypeptide seqid 39285.  
XX  
XX  
XX Plant protectant; plant growth regulator; gene therapy; plant;  
XX recombinant DNA construct; physical array; plant breeding marker;  
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;  
XX extreme osmotic condition; pathogen tolerance; pest tolerance;  
XX growth rate; cell cycle pathway; disease resistance;  
XX galactomannan production; lignin production; plant growth regulator;  
XX yield; plant growth; plant development; seed oil; protein yield;  
XX  
XX protein content.  
XX  
XX Unidentified.  
OS

[illegible]

```

PN      MO200233092-A1.
PD
XX
PE      25-APR-2002.
PF
XX      18-OCT-2001; 2001WO-JP009153.
XX
PR      18-OCT-2000; 2000JP-00318557.
XX
PA      (NAG-) NAT INST AGROBIOLOGICAL SCT.
PI      Yano M, Yamamouchi U;
XX
DR      MPI; 2002-372312/40.
XX
DR      N-PDSB; AAL46958.
XX
PT      Rice-originated gene, Spi17, that inhibits lesion formation and is
XX      applicable in improving heat stress of plants thus leading to prevention
XX      of lesion formation, for developing new breeds of plants for agriculture
XX      and horticulture.
XX
PS      Claim 1; Page 40-42; 53pp; Japanese.
XX
CC      The present invention provides the protein and coding sequences of rice
XX      lesion formation inhibitor Spi17. The protein improves the heat stress of
XX      the plant, and can be used in the development of new breeds of plants for
XX      agriculture and horticulture. The present sequence is the protein of the
XX      invention
SQ      Sequence 459 AA;

Query Match          51.6%; Score 49; DB 5; Length 459;
Best Local Similarity 37.5%; Pred. No. 33;
Matches    6; Conservative   6; Mismatches     4; Indels    0; Gaps    0

QY      1 HLHHVHKEEHAHAN 16
       ||::||:::||::|
DB      102 HLLKNIRKRPVSHS 117

RESULT 31
ABM87520
ID      ABM87520 standard; protein, 459 AA.
XX
AC      ABM87520;
XX
DT      02-JUN-2005 (first entry)
XX
DE      Rice abiotic stress responsive polypeptide SEQ ID NO:5766.
XX
KW      abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.
XX
OS      Oryza sativa.
XX
FN      WO2003008540-A2.
XX
PD      30-JAN-2003.
XX
PF      21-JUN-2002; 2002MO-US019668.
XX
PR      22-JUN-2001; 2001US-0300112P.
XX      PR      24-AUG-2001; 2001US-0314662P.
XX      PR      26-SEP-2001; 2001US-0325277P.
XX      PR      21-NOV-2001; 2001US-0332132P.
XX
PA      (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI      Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,
XX      Moughamer T, Provart N, Ricke D, Zhu T;
XX
DR      MPI; 2003-248011/24.
XX
PT      New stress-responsive nucleic acid, useful for altering the
XX      responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold

```



PT stress, salt stress or osmotic stress.  
 XX  
 PS Claim 1, SEQ ID NO 5766; 89pp; English.  
 XX  
 CC The invention relates to novel abiotic stress responsive polynucleotides  
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host  
 CC cells, and plants containing such polynucleotides. Also disclosed are  
 CC methods for using the polynucleotides and polypeptides to alter the  
 CC responsiveness of a plant to abiotic stress. The invention is useful in  
 CC agriculture. The nucleic acid is useful for determining whether a test  
 CC plant has been exposed to an abiotic stress condition. It is also useful  
 CC for selecting an agent that alters abiotic stress regulated  
 CC polynucleotide expression in a plant cell, and to identify a homolog or  
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid  
 CC molecule and the polypeptide encoded by it are useful in altering the  
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt  
 CC stress, osmotic stress or any of their combinations. The present sequence  
 CC is used in the exemplification of the invention  
 XX  
 SO Sequence 459 AA;  
 Query Match 51.6%; Score 49; DB 7; Length 459;  
 Best Local Similarity 37.5%; Pred. No. 33;  
 Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 HLHNVKKEHAAHN 16  
 DB 102 HLKNIHRRKPVHSHS 117  
 ID ADW17713 standard; protein; 506 AA.  
 XX  
 AC ADW17713;  
 DT 24-MAR-2005 (first entry)  
 XX  
 DE Pinus radiata transcription factor protein HSF family Seq 1480.  
 XX  
 KM plant; transcription; gene regulation; gene expression; transgenic plant;  
 KM drought resistance; disease resistance; salt tolerance; cold tolerance;  
 KM freeing tolerance; flowering; flavor enhancer; flower color.  
 XX  
 OS Pinus radiata.  
 XX  
 PN WO2005001050-A2.  
 XX  
 PD 06-JUN-2005.  
 XX  
 PF 07-JUN-2004; 2004WO-US017965.  
 XX  
 PR 06-JUN-2003; 2003US-0476189P.  
 XX  
 PA (ARBO-) ARBORGEN LLC.  
 XX  
 PI Bloksberg LN, Bryant C, Connell MB, Emerson SJ, Frost MJ,  
 PI Forester RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magnus A;  
 PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;  
 XX  
 DR WPI; 2005-075542/08.  
 DR N-PSDB; ADW16927.  
 XX  
 PT New polynucleotides isolated from plants encoding transcription factors,  
 PT and polypeptides encoded by such polynucleotides, useful for regulating  
 PT gene transcription and gene expression.  
 XX  
 PS Claim 31, SEQ ID NO 1480; 1265pp; English.  
 XX  
 CC This invention relates to novel isolated plant nucleic acid molecules, or  
 CC variants thereof, that encode transcription factors. Specifically, it  
 CC refers to transcription factor proteins that are capable of binding to  
 CC DNA in order to regulate gene transcription and gene expression in a

CC plants, in particular Eucalyptus grandis and Pinus radiata. The present  
 CC invention describes DNA constructs containing DNA encoding a  
 CC transcription factor that regulates the promoter, which is operably  
 CC linked to the desired nucleic acid to be expressed. It further provides  
 CC transgenic plants expressing a transcription factor that confers a trait  
 CC to the plant such as increased drought, salt or disease tolerance, height  
 CC change, enhanced cold/frost tolerance, enhanced color, health and  
 CC nutritional characteristics, as well as improved taste, starch  
 CC composition, flower longevity and germination, amongst others.  
 CC Accordingly, such plants that are successfully transfected with a DNA  
 CC construct can be characterized by a difference in flower color, petal or  
 CC leaf shape and size, aroma or plant height. This polypeptide is a plant  
 CC transcription factor protein sequence of the invention.  
 XX  
 SO Sequence 506 AA;  
 Query Match 51.6%; Score 49; DB 9; Length 506;  
 Best Local Similarity 37.5%; Pred. No. 38;  
 Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 HLHNVKKEHAAHN 16  
 DB 96 HLKNIHRRKPVHSHS 111  
 ID AAW99482 standard; protein; 1463 AA.  
 XX  
 AC AAW99482;  
 DT 08-JUN-1999 (first entry)  
 XX  
 DE Murine NCoA-2 protein.  
 XX  
 KM Mouse; p/CTP; p300/CBP/co-integrator-associated protein; gene expression;  
 KM breast cancer; inflammatory disease; atherosclerosis; osteoporosis.  
 XX  
 OS Mus sp.  
 XX  
 OS  
 XX  
 FT Key Location/Qualifiers  
 FT Misc-difference 989  
 FT /label= unknown  
 XX  
 PN WO9856806-A1.  
 XX  
 PD 17-DEC-1998.  
 XX  
 PF 12-JUN-1998; 98WO-US012263.  
 XX  
 PR 12-JUN-1997; 97US-0049452P.  
 XX  
 PA (REGC ) UNITV CALIFORNIA.  
 XX  
 PI Rosenfield MG, Glass CK, Rose DW, Torchia J;  
 XX  
 DR WPI; 1999-080863/07.  
 XX  
 PT New nucleic acids encoding p/CTP and NCoA-2 polypeptides - are used to  
 PT identify agents that regulate gene expression, e.g. for treatment of  
 PT cancer, inflammatory disease and osteoporosis.  
 XX  
 PS Claim 22; Fig 2A; 100pp; English.  
 XX  
 CC This sequence represents the amino acid sequence of the mouse NCoA-2. The  
 CC protein can regulate gene expression so are potentially useful  
 CC therapeutically, e.g. against (breast) cancer, inflammatory disease (e.g.  
 CC atherosclerosis) or osteoporosis  
 XX  
 SO Sequence 1463 AA;  
 Query Match 51.6%; Score 49; DB 2; Length 1463;  
 Best Local Similarity 53.3%; Pred. No. 1.3e+02;

Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 HLIHNVKKEHAAH 15  
 DB 424 HEPHNGPKQHGPH 438

RESULT 34  
 ID ABP07487 standard; protein; 74 AA.  
 AC ABP07487;  
 DT 25-JUN-2002 (first entry)  
 DE Human ORFX protein sequence SEQ ID NO:14956.  
 XX  
 XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
 KM hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
 KM degenerative disorder; osteoarthritis; neurodegenerative disorder;  
 KM cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
 KM hypertension; hypothyroidism; cholesterol ester storage disease;  
 KM immune deficiency; immune disorder; infectious disease;  
 KM autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
 KM myasthenia gravis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200192523-A2.  
 PD 06-DEC-2001.  
 PF 29-MAY-2001; 2001WO-US010836.  
 PR 30-MAY-2000; 2000US-0206132P.  
 PR 29-AUG-2000; 2000US-0228716P.  
 XX  
 PA (CUNA-) CUNAGEN CORP.  
 PI Shinketsu RA, Leach MD;  
 DR WPI; 2002-106308/14.  
 DR N-PSDB; ABN23239.  
 XX  
 PT Novel human polypeptides and polynucleotides useful for diagnosing,  
 PT preventing and treating cardiovascular disease, neurodegenerative,  
 PT hyperproliferative disorders and autoimmune disorders.  
 PT  
 PS Disclosure; SEQ ID NO 14956; 1037pp; English.  
 XX  
 XX The present invention describes substantially purified human proteins  
 CC (referred to as open reading frame, ORFX, where X is 1-1191 (see table 1  
 CC in the specification). ABN15762 to ABN27232 encode the human ORFX  
 CC proteins given in ABP00010 to ABP1500. ORFX proteins are useful for  
 CC creating or preventing a pathology associated with an ORFX-associated  
 CC disorder in humans, and in the manufacture of a medicament for treating a  
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
 CC sequences can be used in gene therapy. ORFX sequences can be used in the  
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
 CC storage disease, various immune deficiencies and disorders, infectious  
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
 CC bone degenerative disorders, or periodontal disease, and for gut  
 CC protection or regeneration and treatment of lung or liver fibrosis,  
 CC reperfusion injury in various tissues and conditions resulting from  
 CC systemic cytokine damage. N.B. The sequence data for this patent did not  
 CC form part of the printed specification, but was obtained in electronic

CC format directly from WIPO at ftp.wipo.int/pub/published\_pat\_sequences  
 XX  
 SQ Sequence 74 AA;  
 Query Match 50.5%; Score 48; DB 5; Length 74;  
 Best Local Similarity 43.8%; Pred. No. 5.5;  
 Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 HLIHNVKKEHAAH 16  
 DB 8 HHTHRTTHHAAHH 23

RESULT 35  
 ID ADN73121 standard; protein; 86 AA.  
 AC ADN73121;  
 DT 15-JUL-2004 (first entry)  
 DE Thale cress protein upregulated in E2Fa/Dpa expressing plants SegID 1016.  
 XX  
 KM plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;  
 KM animal feed product; thale cress; cell wall biosynthesis;  
 KM nitrogen metabolism; carbon metabolism.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN WO2004035798-A2.  
 PD 29-APR-2004.  
 PF 20-OCT-2003; 2003WO-EP011658.  
 PR 18-OCT-2002; 2002EP-00079408.  
 XX  
 PA (CROP-) CROPDESIGN NV.  
 PI Inze D, De Veylder L, Vlieghe K;  
 DR WPI; 2004-348466/32.  
 DR N-PSDB; ADN73120.  
 XX  
 PT Altering plant characteristics, useful for producing plants for enzyme or  
 PT pharmaceutical production comprises modifying in a plant, expression of  
 PT one or more nucleic acids and/or modifying level or activity of one or  
 PT more proteins.  
 PT  
 PS Claim 1; SEQ ID NO 1016; 134pp; English.  
 XX  
 XX This invention relates to a novel method for altering one or more plant  
 CC characteristics. Specifically, it refers to identifying genes that are up  
 CC - or down-regulated in transgenic plants overexpressing the heterodimeric  
 CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to  
 CC alter plant characteristics accordingly. The present invention describes  
 CC generating transgenic plants for the production of growth regulators,  
 CC enzymes, therapeutics, pharmaceuticals and animal feed products, where  
 CC the altered plant characteristics are selected from increased yield or  
 CC biomass, enhanced survival capacity, stress tolerance, plant architecture  
 CC or physiology, altered endoreplication, biochemistry, signal  
 CC transduction, storage lipid mobilisation and/or altered photosynthesis,  
 CC each relative to the corresponding wild type plants. Accordingly, these  
 CC sequences can also be useful as positive or negative selectable markers  
 CC during transformation of cells or tissues. The identified genes play a  
 CC role in a variety of biological processes such as DNA replication, cell  
 CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as  
 CC transcription factors. This polypeptide sequence is thale cress protein  
 CC expressed by a gene upregulated 1.3 fold or more in plants overexpressing  
 CC the E2Fa/Dpa transcription factor, given in an exemplification of the  
 CC invention.  
 XX  
 SQ Sequence 86 AA;

Query Match 50.5%; Score 48; DB 8; Length 86;  
Best Local Similarity 40.0%; Pred. No. 6;  
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;  
QY 1 HLHNVHKEEHAHAH 15  
|:|:|:|:|:|:|:  
Db 70 HHHHHNEEHCICY 84

RESULT 36  
ABO62756  
ID ABO62756 standard; protein; 161 AA.  
XX  
AC ABO62756;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Klebsiella pneumoniae polypeptide seqid 9273.  
XX  
KW Recombinant expression vector; transcription regulatory element;  
KW Klebsiella pneumoniae protein; antibacterial; Vaccine.  
XX  
OS Klebsiella pneumoniae.  
XX  
PN US6610836-B1.  
XX  
PD 26-AUG-2003.  
XX  
PF 27-JAN-2000; 2000US-00489039.  
XX  
PR 29-JAN-1999; 99US-0117747P.  
XX  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Breton GL, Osborne M;  
XX  
DR WPI, 2003-895346/82.  
XX  
DR N-PSDB; ACH96307.  
XX  
PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for  
PT preparing a vaccine composition against Klebsiella pneumoniae.  
XX  
XX  
PS Disclosure, SEQ ID NO 9273; 932PP; English.  
XX  
CC The invention describes a new isolated nucleic acid encoding a Klebsiella  
CC pneumoniae polypeptide. Also described are: a recombinant expression  
CC vector comprising the nucleic acid, operably linked to a transcription  
CC regulatory element; and a cell comprising the recombinant expression  
CC vector. The nucleic acid is useful for preparing a vaccine composition  
CC against Klebsiella pneumoniae. This is the amino acid sequence of a  
CC Klebsiella pneumoniae polypeptide of the invention  
XX  
SQ Sequence 161 AA;

Query Match 50.5%; Score 48; DB 7; Length 161;  
Best Local Similarity 46.7%; Pred. No. 14;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
QY 1 HLHNVHKEEHAHAH 15  
|:|:|:|:|:|:|:  
Db 147 HHHHHHHHHHHSH 161

RESULT 37  
ADSI17428  
ID ADSI17428 standard; protein; 191 AA.  
XX  
AC ADSI17428;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Amino acid sequence of a heat shock transcription factor.

XX  
KW phyto remediation; cadmium/zinc transporting P-type ATPase;  
KW Cd/Zn transporting P-type ATPase; Cpx-ATPase; soil contamination;  
KW heavy metal; cadmium; phytoextraction; cadmium tolerance;  
KW metal detoxification; heavy metal tolerance; metal binding protein;  
KW metallochionein; phytochelatin; TcHMA4; plant;  
KW heat shock transcription factor.  
XX  
OS Thlaapi caerulescens.  
XX  
PN WO2004078905-A2.  
XX  
PD 16-SEP-2004.  
XX  
PF 08-MAR-2004; 2004WO-BE000035.  
XX  
PR 07-MAR-2003; 2003US-0453271P.  
XX  
XX (ULBR ) UNIV LIBRE BRUXELLES.  
XX  
PI Verbruggen N, Bernard C;  
XX  
DR WPI; 2004-662412/64.  
XX  
DR N-PSDB; ADSI17427.  
XX  
PT Novel isolated and purified polypeptide having sequence comprising  
PT fragment of potential cadmium/zinc transporting P-type ATPase, useful in  
PT phyto remediation treatment of medium, preferably soil, contaminated by  
XX heavy metals e.g. cadmium.  
XX  
PS Disclosure; SEQ ID NO 14; 46PP; English.  
XX  
XX The specification describes polypeptides which are useful in  
XX phyto remediation. These polypeptides have sequence identify with a  
XX potential cadmium/zinc (Cd/Zn) transporting P-type ATPase (see ADSI17418)  
XX (Cpx-ATPase). Plants transformed with phyto remediation polypeptides of  
XX the invention are useful for phyto remediation treatment of a medium, e.g.  
XX soil, contaminated by heavy metals such as cadmium, and for  
XX phytoextraction applications. Such plants have improved cadmium  
XX tolerance. The present sequence represents a heat shock transcription  
XX factor. It was isolated from a cDNA library of Thlaapi caerulescens which  
XX was screened in Saccharomyces cerevisiae growing on medium supplemented  
XX with cadmium. Plasmids were rescued from transformants growing on the  
XX cadmium-supplemented medium, and used to re-transform yeast which was  
XX then retested by growing on cadmium-supplemented medium. From sequence  
XX analysis of resulting clones, 19 different non-redundant cDNAs were  
XX identified. The majority of the identified cDNAs encode proteins known to  
XX have a potential role in heavy metal tolerance as metal binding proteins,  
XX metallochioneins and phytochelatins, and putative Cpx-ATPases. As the  
XX identified cDNAs encoding truncated putative Cpx-ATPases showed a higher  
XX similarity with the C-terminus of Arabidopsis thaliana HMA4, these  
XX corresponding sequences in T. caerulescens were referred as TcHMA4.

Query Match 50.5%; Score 48; DB 8; Length 191;  
Best Local Similarity 37.5%; Pred. No. 17;  
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;  
QY 1 HLHNVHKEEHAHAH 16  
|:|:|:|:|:|:|:  
Db 97 HLMKNHRRKPVHSHS 112

RESULT 38  
ADP05866  
ID ADP05866 standard; protein; 400 AA.  
XX  
AC ADP05866;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Bacterial polypeptide #1979.

XX	25-MAY-2001.
PD	
XX	
PF	14-NOV-2000; 2000MO-US031418.
XX	
PR	17-NOV-1999; 99US-0166228P.
PR	17-APR-2000; 2000US-0197899P.
BR	22-AUG-2000; 2000US-0227439P.
XX	
PA	(MEND-) MENDEL BIOTECHNOLOGY INC.
PA	(HEAR/) HEARD J.
PA	(RATC/) RATCLIFFE O.
PA	(GREE/) CREELMAN R.
PA	(JIANG/) JIANG C.
PA	(PINE/) PINEDA O.
PA	(REUB/) REUBER L.
PA	(ADAM/) ADAM L.
XX	
PI	Heard J, Ratcliffe O, Creelman R, Jiang C, Pineda O, Reuber L,
PI	Adam L,
XX	
DR	WPI; 2001-335978/35.
DR	N-PSDB; AAB06474.
XX	
PT	Nucleic acids encoding plant transcription factor polypeptides, useful
PT	for altering the pathogen resistance characteristics of plants, e.g.
PT	corn, potato and cotton plants.
XX	
PS	Claim 4; Page 65-66; 134pp; English.
XX	
CC	The present sequence is Arabidopsis thaliana transcription factor. The
CC	transcription factors are used to alter the structure and developmental
CC	characteristics of plants such as soybean, wheat, corn, potato, cotton,
CC	rice, oilseed rape, sunflower, alfalfa, sugarcane, turf, banana,
CC	blackberry, blueberry, strawberry, raspberry, carrot, cantaloupe,
CC	calliflower, coffee, cucumber, eggplant, grapes, mango, lettuce,
CC	honeydew, melon, onion, papaya, pear, peppers, pineapple, spinach,
CC	squash, sweet corn, tobacco, tomato, watermelon, rosaceous fruits and
CC	vegetable brassicas. The transcription factors are specifically useful
CC	for modifying traits associated with plant's pathogen tolerance such as
CC	alterations in cell wall composition, trichome number or structure,
CC	callose induction, phytoalexin induction, and alterations in the cell
CC	death response. Transgenic plants expressing these transcription factors
CC	are more tolerant to biotrophic or necrotrophic pathogens such as fungi,
CC	bacteria, molluscs, viruses, nematodes and parasitic higher plants. The
CC	transcription factors are also used in gene therapy
XX	
SQ	Sequence 401 AA;
XX	
Query Match	50.5%; Score 48; DB 4; Length 401;
Best Local Similarity	37.5%; Pred. No. 41;
Matches	6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
OY	1 HLINHYKESEHAAHN 16
	: : : : : : : :
DB	97 HLMKNHRRKPVSHS 112
XX	
RESULT 40	
ABO43132	
ID	ABO43132 standard; protein; 401 AA.
XX	
AC	ABO43132;
XX	
DT	23-SEP-2003 (first entry)
XX	
DE	A. thaliana disease tolerance transcription factor, G261.
XX	
KM	Plant; transcription factor; disease resistance; transgenic;
KM	plant breeding; pathogens resistance; pests; resistance.
XX	
OS	Arabidopsis thaliana.
XX	

Search completed: February 11, 2006, 13:10:40  
Job time : 203 secs

PN US2003046723-A1.  
XX  
PD 06-MAR-2003.  
XX  
XX 22-MAR-2000; 2000US-00533029.  
PF  
XX 22-MAR-2000; 2000US-00533029.  
PR  
XX  
PA (HEAR/) HEARD J.  
PA (BROU/) BROUN P.  
PA (RIEC/) RIECHMANN J L.  
PA (KEDD/) KEDDIE J.  
PA (PINE/) PINEDA O.  
PA (ADAM/) ADAM L.  
PA (SAMA/) SAMAH R.  
PA (ZHAN/) ZHANG J.  
PA (YUGG/) YU G.  
PA (RATC/) RATCLIFFE O.  
PA (PIIG/) PILGRIM M.  
PA (JIANG/) JIANG C.  
PA (REUB/) REUBER L.  
XX  
XX  
PI Heard J, Brown P, Riechmann JL, Keddie J, Pineda O, Adam L,  
PI Samaha R, Zhang J, Yu G, Ratcliffe O, Pilgrim M, Jiang C, Reuber L,  
XX  
XX WPI; 2003-521768/49.  
DR N-PSDB; ACD98398.  
XX  
XX  
PT New transgenic plants comprising a recombinant gene that alters the  
PT plant's disease tolerance or resistance, useful in plant breeding, e.g.  
PT for generating plants with improved tolerance or resistance to diseases,  
PT pests or pathogens.  
XX  
XX  
PS Claim 1; Page 78-79; 124pp; English.  
XX  
XX  
CC The invention relates to a transgenic plant, comprising a recombinant  
CC polynucleotide that alters the plant's disease tolerance or resistance  
CC when compared with the same trait of another plant lacking the  
CC recombinant polynucleotide. The recombinant polynucleotide comprises a  
CC nucleotide sequence, which encodes a polypeptide comprising at least 6  
CC consecutive amino acids of any of 56 transcription factor proteins  
CC appearing as AB043093-AB043148. Also included are altering the disease  
CC tolerance or resistance of a plant (b); (a) transforming a plant with the  
CC recombinant polynucleotide; (b) selecting the transformed plants; and (c)  
CC identifying a transformed plant with an altered disease tolerance or  
CC resistance), altering the expression levels of at least one gene in a  
CC plant by transforming the plant with the recombinant polynucleotide,  
CC altering a plant's trait (comprising: (a) providing a database sequence;  
CC (b) comparing the database sequence with the polypeptide or  
CC polynucleotide cited above; (c) selecting a database sequence that meets  
CC the selected sequence criteria; and (d) transforming the selected  
CC database sequence in the plant) and altering a plant's trait (comprising:  
CC (a) providing a test polynucleotide; (b) hybridizing the test  
CC polynucleotide at low stringency with the recombinant polynucleotide  
CC cited above; and (c) transforming the hybridizing test polynucleotide in  
CC a plant to alter a trait of the plant. The transgenic plant is useful in  
CC plant breeding, particularly for generating plants with improved  
CC tolerance or resistance to diseases. The plants have commercial utility  
CC for increasing tolerance or resistance to pathogens and pests. The  
CC present sequence is an Arabidopsis thaliana transcription factor of the  
CC invention  
XX  
SQ Sequence 401 AA;  
XX  
Query Match 50.5%; Score 48; DB 7; Length 401;  
Best Local Similarity 37.5%; Pred. No. 41;  
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;  
QY 1 HLINHVKEEHAHNN 16  
Db 97 HLMKNITHRRKPVHSHS 112

"This Page Blank (copy)"

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2006, 13:07:22 ; Search time 251 Seconds  
(without alignments)  
44.974 Million cell updates/sec

Title: US-10-762-588-1

Perfect score: 95  
Sequence: 1 HLHNHVKESHMAHNN 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : UniProt\_05.80:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95	100.0	331	1	LDHA_CHICK
2	91	95.8	331	1	LDHA_ALUMI
3	91	95.8	331	1	LDHA_CAICA
4	84	88.4	331	1	LDHA_COLLI
5	76	80.0	331	1	LDHA_PELSI
6	76	80.0	331	1	LDHA_TRASC
7	76	80.0	332	2	O6S5M1_MACTB
8	76	80.0	332	2	O6S5M4_9SAUR
9	76	80.0	332	2	O6S5M6_APAPE
10	73	76.8	332	2	O6S5M2_9SAUR
11	72	75.8	332	2	O6YL22_IGUIG
12	66	69.5	331	1	LDHA_PYRG
13	66	69.5	331	1	LDHA_PYRG
14	62	65.3	331	1	LDHA_SCEUN
15	59	62.1	662	2	O9LTX1_ARATH
16	59	62.1	674	2	O9C7B4_ARATH
17	54.5	57.4	163	2	O5E729_VIBF1
18	54	56.8	1032	2	O7SCV2_NEUCR
19	53	55.8	725	2	O8T990_DROME
20	53	55.8	837	2	O9W4B2_DROME
21	53	55.8	837	2	O5UID9_DROME
22	52	54.7	717	2	O6C9B8_YARLI
23	52	54.7	1018	2	O7S3X5_NEUCR
24	51	53.7	339	2	O8INR1_DROME
25	51	53.7	402	1	POXN_DROME
26	51	53.7	2126	2	O8IDA6_PLAT7
27	50	52.6	245	2	O9FUB1_ARATH
28	50	52.6	326	2	O6G7G4_STAAS
29	50	52.6	326	2	O8NVF2_STAAM
30	50	52.6	351	1	HRPX_PLALO
31	50	52.6	448	2	O9LWRL_ARATH

32	50	52.6	735	2	O9NE57_CABEL	O9NE57_CABEL
33	49.5	52.1	498	2	O6VSC5_ORYSA	O6VSC5_ORYSA
34	49	51.6	261	2	O8U226_9EUDU	O8U226_9EUDU
35	49	51.6	374	2	O6NLA8_DROME	O6NLA8_DROME
36	49	51.6	374	2	O9VFN8_DROME	O9VFN8_DROME
37	49	51.6	408	2	O9SXX8_TOBAC	O9SXX8_TOBAC
38	49	51.6	417	2	O7SP82_ASHCO	O7SP82_ASHCO
39	49	51.6	411	2	O8T880_ASPNO	O8T880_ASPNO
40	49	51.6	442	2	O5VD75_ASPNO	O5VD75_ASPNO
41	49	51.6	442	2	O8TG77_ASPNO	O8TG77_ASPNO
42	49	51.6	442	2	O8TG78_ASPNO	O8TG78_ASPNO
43	49	51.6	442	2	O8TG79_ASPNO	O8TG79_ASPNO
44	49	51.6	442	2	O8TG81_ASPNO	O8TG81_ASPNO
45	49	51.6	442	2	O8TG82_ASPNO	O8TG82_ASPNO
46	49	51.6	442	2	O8TG83_ASPNO	O8TG83_ASPNO
47	49	51.6	442	2	O8TG84_ASPNO	O8TG84_ASPNO
48	49	51.6	442	2	O8TG85_ASPNO	O8TG85_ASPNO
49	49	51.6	442	2	O8TG86_ASPNO	O8TG86_ASPNO
50	49	51.6	459	2	O93VB5_ORYSA	O93VB5_ORYSA
51	49	51.6	459	2	O93WD1_ORYSA	O93WD1_ORYSA
52	49	51.6	572	2	O6FJ90_CANGA	O6FJ90_CANGA
53	49	51.6	1642	2	O5AN92_DICDI	O5AN92_DICDI
54	48.5	51.1	250	2	O6AN47_DESPS	O6AN47_DESPS
55	48	50.5	86	2	O9W9M3_ARATH	O9W9M3_ARATH
56	48	50.5	158	1	UREB_KTEAR	UREB_KTEAR
57	48	50.5	183	2	O8WZ20_HELNU	O8WZ20_HELNU
58	48	50.5	256	2	O4NMW6_BACCE	O4NMW6_BACCE
59	48	50.5	256	2	O634Q6_BACCE	O634Q6_BACCE
60	48	50.5	256	2	O6HND6_BACCK	O6HND6_BACCK
61	48	50.5	256	2	O730Q5_BACCI	O730Q5_BACCI
62	48	50.5	256	2	O818H5_BACCR	O818H5_BACCR
63	48	50.5	256	2	O81LV4_BACNA	O81LV4_BACNA
64	48	50.5	275	2	O6ZPH0_ORYSA	O6ZPH0_ORYSA
65	48	50.5	401	2	O82078_ARATH	O82078_ARATH
66	48	50.5	401	2	O49403_ARATH	O49403_ARATH
67	48	50.5	402	2	O940V6_PHAAT	O940V6_PHAAT
68	48	50.5	402	2	O9MS97_MEDSA	O9MS97_MEDSA
69	48	50.5	418	2	O7QZ17_GIALA	O7QZ17_GIALA
70	48	50.5	466	2	O94AB25_ARATH	O94AB25_ARATH
71	48	50.5	475	2	O6K6S5_ORYSA	O6K6S5_ORYSA
72	48	50.5	652	2	O9NS42_HUMAN	O9NS42_HUMAN
73	48	50.5	706	1	FOI_DROME	FOI_DROME
74	48	50.5	834	2	O23259_ARATH	O23259_ARATH
75	48	50.5	1078	2	O5CVD2_CRYPO	O5CVD2_CRYPO
76	48	50.5	1164	2	O4IBI1_GIBZE	O4IBI1_GIBZE
77	48	50.5	1235	2	O7R920_PLAYO	O7R920_PLAYO
78	48	50.5	1266	2	O8NIV2_NEUCR	O8NIV2_NEUCR
79	48	50.5	1292	2	O20330_CABEL	O20330_CABEL
80	47	49.5	115	2	O7YV22_9TRYP	O7YV22_9TRYP
81	47	49.5	142	2	O8TBF1_PLAFA	O8TBF1_PLAFA
82	47	49.5	194	2	O5FH84_EHRRG	O5FH84_EHRRG
83	47	49.5	194	2	O5HB25_EHRRW	O5HB25_EHRRW
84	47	49.5	199	2	O61IT5_DROME	O61IT5_DROME
85	47	49.5	317	2	O6MB36_PARUM	O6MB36_PARUM
86	47	49.5	388	2	O62BH7_CABER	O62BH7_CABER
87	47	49.5	399	2	O8MSK2_DROME	O8MSK2_DROME
88	47	49.5	464	2	O9W249_DROME	O9W249_DROME
89	47	49.5	520	2	O8XOT1_NEUCR	O8XOT1_NEUCR
90	47	49.5	591	2	O5AM41_CANAL	O5AM41_CANAL
91	47	49.5	593	2	O5ALP2_CANAL	O5ALP2_CANAL
92	47	49.5	639	2	O5AG32_CANAL	O5AG32_CANAL
93	47	49.5	639	2	O5AG37_CANAL	O5AG37_CANAL
94	47	49.5	645	1	BABH2_DROME	BABH2_DROME
95	47	49.5	661	2	O84EK7_SHEFR	O84EK7_SHEFR
96	47	49.5	697	2	O6AMM4_DROME	O6AMM4_DROME
97	47	49.5	697	2	O81GT7_DROME	O81GT7_DROME
98	47	49.5	697	2	O81MV7_DROME	O81MV7_DROME
99	47	49.5	697	2	O9VC50_DROME	O9VC50_DROME
100	47	49.5	710	2	O4Y806_PLACH	O4Y806_PLACH

## ALIGNMENTS

O9NE57_CABEL	O9NE57_CABEL
O6VSC5_ORYSA	O6VSC5_ORYSA
O8U226_9EUDU	O8U226_9EUDU
O6NLA8_DROME	O6NLA8_DROME
O9VFN8_DROME	O9VFN8_DROME
O9SXX8_TOBAC	O9SXX8_TOBAC
O7SP82_ASHCO	O7SP82_ASHCO
O8T880_ASPNO	O8T880_ASPNO
O5VD75_ASPNO	O5VD75_ASPNO
O8TG77_ASPNO	O8TG77_ASPNO
O8TG78_ASPNO	O8TG78_ASPNO
O8TG79_ASPNO	O8TG79_ASPNO
O8TG81_ASPNO	O8TG81_ASPNO
O8TG82_ASPNO	O8TG82_ASPNO
O8TG83_ASPNO	O8TG83_ASPNO
O8TG84_ASPNO	O8TG84_ASPNO
O8TG85_ASPNO	O8TG85_ASPNO
O8TG86_ASPNO	O8TG86_ASPNO
O93VB5_ORYSA	O93VB5_ORYSA
O93WD1_ORYSA	O93WD1_ORYSA
O6FJ90_CANGA	O6FJ90_CANGA
O5AN92_DICDI	O5AN92_DICDI
O6AN47_DESPS	O6AN47_DESPS
O9W9M3_ARATH	O9W9M3_ARATH
UREB_KTEAR	UREB_KTEAR
O8WZ20_HELNU	O8WZ20_HELNU
O4NMW6_BACCE	O4NMW6_BACCE
O634Q6_BACCE	O634Q6_BACCE
O6HND6_BACCK	O6HND6_BACCK
O730Q5_BACCI	O730Q5_BACCI
O818H5_BACCR	O818H5_BACCR
O81LV4_BACNA	O81LV4_BACNA
O6ZPH0_ORYSA	O6ZPH0_ORYSA
O82078_ARATH	O82078_ARATH
O49403_ARATH	O49403_ARATH
O940V6_PHAAT	O940V6_PHAAT
O9MS97_MEDSA	O9MS97_MEDSA
O7QZ17_GIALA	O7QZ17_GIALA
O94AB25_ARATH	O94AB25_ARATH
O6K6S5_ORYSA	O6K6S5_ORYSA
O9NS42_HUMAN	O9NS42_HUMAN
FOI_DROME	FOI_DROME
O23259_ARATH	O23259_ARATH
O5CVD2_CRYPO	O5CVD2_CRYPO
O4IBI1_GIBZE	O4IBI1_GIBZE
O7R920_PLAYO	O7R920_PLAYO
O8NIV2_NEUCR	O8NIV2_NEUCR
O20330_CABEL	O20330_CABEL
O7YV22_9TRYP	O7YV22_9TRYP
O8TBF1_PLAFA	O8TBF1_PLAFA
O5FH84_EHRRG	O5FH84_EHRRG
O5HB25_EHRRW	O5HB25_EHRRW
O61IT5_DROME	O61IT5_DROME
O6MB36_PARUM	O6MB36_PARUM
O62BH7_CABER	O62BH7_CABER
O8MSK2_DROME	O8MSK2_DROME
O9W249_DROME	O9W249_DROME
O8XOT1_NEUCR	O8XOT1_NEUCR
O5AM41_CANAL	O5AM41_CANAL
O5ALP2_CANAL	O5ALP2_CANAL
O5AG32_CANAL	O5AG32_CANAL
O5AG37_CANAL	O5AG37_CANAL
BABH2_DROME	BABH2_DROME
O84EK7_SHEFR	O84EK7_SHEFR
O6AMM4_DROME	O6AMM4_DROME
O81GT7_DROME	O81GT7_DROME
O81MV7_DROME	O81MV7_DROME
O9VC50_DROME	O9VC50_DROME
O4Y806_PLACH	O4Y806_PLACH



```

RESULT 1
ID  LDHA_CHICK      STANDARD;      PRT;      331 AA.
AC  P00340;
DT  21-JUL-1986 (Rel. 01, Created)
DT  01-NOV-1991 (Rel. 20, Last sequence update)
DT  10-MAY-2005 (Rel. 47, Last annotation update)
DE  L-lactate dehydrogenase A chain (EC 1.1.1.27) (LDH-A).
GN  Name-LDHA;
OS  Gallus gallus (Chicken).
OC  Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC  Gallus.
OX  NCBI_TaxID=9031;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RX  MEDLINE=91057138; PubMed=2243792;
RA  Hirota Y., Katsumata A., Takeya T.;
RT  "Nucleotide and deduced amino acid sequences of chicken lactate
RT  dehydrogenase-A."
RL  Nucleic Acids Res. 18:6432-6432(1990).
RN  [2]
RP  PROTEIN SEQUENCE.
RA  Torff H.-J., Becker D., Schwarzwald J.;
RL  (in) Sund H. (eds.);
RT  Pyridine nucleotide dependent dehydrogenases, pp.31-42, Walter de
RL  Gruyter, Berlin (1977).
CC  -1- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) = pyruvate + NADH.
CC  -1- PATHWAY: Anaerobic glycolysis; final step.
CC  -1- SUBUNIT: Homotetramer.
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -1- SIMILARITY: Belongs to the LDH/MDH superfamily. LDH family.
CC
CC  This Swiss-Prot entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use as long as its content is in no way modified and this statement is not
CC  removed.
CC
CC  -----
CC  EMBL; X53828; CAA37824.1; -; mRNA.
CC
CC  DR  PIR; A00349; DECHLM.
CC  DR  PIR; S12151; S12151.
CC  DR  HSSP; P00338; 1110.
CC  DR  SMR; P00340; 1-331.
CC  DR  ENSEMBL; ENSGALG0000006300; Gallus gallus.
CC  DR  InterPro; IPR011304; L-LDH-NAD.
CC  DR  InterPro; IPR001557; L-LDH_MDH.
CC  DR  InterPro; IPR001236; Ldh.
CC  DR  PANTHER; PTHR11540; Ldh; 1.
CC  DR  Pfam; PF02866; Ldh_1_C; 1.
CC  DR  Pfam; PF00056; Ldh_1_N; 1.
CC  DR  PRINTS; PIRSF00102; Lac_mal_DH; 1.
CC  DR  PRINTS; PRO0086; LLDHGRNAS.
CC  DR  TIGRFS; TIGR01771; L-LDH-NAD; 1.
CC  DR  PROSITE; PS00064; L-LDH; 1.
CC  DR  Direct protein sequencing; Glycolysis; Multigene family; NAD;
CC  Oxidoreductase.
CC  KM
CC  FT  INIT MET      0
CC  FT  NP BIND      28      56      NAD (By similarity).
CC  FT  ACT SITE     192     192      Proton acceptor (By similarity).
CC  FT  BINDING      98      98      NAD (By similarity).
CC  FT  BINDING     105     105      Substrate (By similarity).
CC  FT  BINDING     137     137      NAD or substrate (By similarity).
CC  FT  BINDING     161     168      Substrate (By similarity).
CC  FT  BINDING     247     247      Substrate (By similarity).
CC  FT  BINDING      62      62      L -> M (in Ref. 2).
CC  FT  CONFLICT     77      77      I -> T (in Ref. 2).
CC  FT  CONFLICT     191     191      E -> Q (in Ref. 2).
CC  FT  CONFLICT     191     191
CC  SQ  SEQUENCE      331 AA; 36383 MW; FDD94D8C90A3DF1 CRC64;

```

Query Match 100.0%; Score 95; DB 1; Length 331;  
 Best Local Similarity 100.0%; Pred. No. 4e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLIHNVKKEHNAHN 16  
 |||||  
 DB 5 HLIHNVKKEHNAHN 20

```

RESULT 2
ID  LDHA_ALIMI      STANDARD;      PRT;      331 AA.
AC  Q9PW06;
DT  28-FEB-2003 (Rel. 41, Created)
DT  28-FEB-2003 (Rel. 41, Last sequence update)
DT  10-MAY-2005 (Rel. 47, Last annotation update)
DE  L-lactate dehydrogenase A chain (EC 1.1.1.27) (LDH-A).
GN  Name-LDHA;
OS  Alligator mississippiensis (American alligator).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX  NCBI_TaxID=8496;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  TISSUE=Muscle;
RX  MEDLINE=98031373; PubMed=9364765;
RA  Mannan H., Teol S.C.-M., Krishnak J.S., Li W.-H., Li S.S.-L.;
RT  "The cDNA cloning and molecular evolution of reptile and pigeon
RT  lactate dehydrogenase isozymes."
RL  Mol. Biol. Evol. 14:1081-1087(1997).
CC  -1- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) = pyruvate + NADH.
CC  -1- PATHWAY: Anaerobic glycolysis; final step.
CC  -1- SUBUNIT: Homotetramer (By similarity).
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC  -1- SIMILARITY: Belongs to the LDH/MDH superfamily. LDH family.
CC
CC  This Swiss-Prot entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use as long as its content is in no way modified and this statement is not
CC  removed.
CC
CC  -----
CC  EMBL; U79511; AAD46977.1; -; mRNA.
CC
CC  DR  HSSP; P00339; 9LDT.
CC  DR  InterPro; IPR011304; L-LDH-NAD.
CC  DR  InterPro; IPR001557; L-LDH_MDH.
CC  DR  InterPro; IPR001236; Ldh.
CC  DR  PANTHER; PTHR11540; Ldh; 1.
CC  DR  Pfam; PF02866; Ldh_1_C; 1.
CC  DR  Pfam; PF00056; Ldh_1_N; 1.
CC  DR  PRINTS; PIRSF00102; Lac_mal_DH; 1.
CC  DR  PRINTS; PRO0086; LLDHGRNAS.
CC  DR  TIGRFS; TIGR01771; L-LDH-NAD; 1.
CC  DR  PROSITE; PS00064; L-LDH; 1.
CC  DR  Glycolysis; Multigene family; NAD; Oxidoreductase.
CC  KM
CC  FT  INIT MET      0
CC  FT  NP BIND      28      56      NAD (By similarity).
CC  FT  ACT SITE     192     192      Proton acceptor (By similarity).
CC  FT  BINDING      98      98      NAD (By similarity).
CC  FT  BINDING     105     105      Substrate (By similarity).
CC  FT  BINDING     137     137      NAD or substrate (By similarity).
CC  FT  BINDING     168     168      Substrate (By similarity).
CC  FT  BINDING     247     247      Substrate (By similarity).
CC  SQ  SEQUENCE      331 AA; 36652 MW; 5863F9CDD21A192 CRC64;

```

Query Match 95.8%; Score 91; DB 1; Length 331;  
 Best Local Similarity 93.8%; Pred. No. 1.6e-05;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLIHNVKKEHNAHN 16  
 |||||  
 DB 5 HLIHNVKKEHNAHN 20

RESULT 3

DE l-lactate dehydrogenase A chain (EC 1.1.1.27) (LDH-A).  
GN Name=LDHA;  
OS Columba livia (Domestic pigeon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Archaeosuria; Aves; Neognathae; Columbiformes; Columbidae; Columba.  
CX NCBI\_TaxID=6932;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Muscle;  
RX MEDLINE=96031373; PubMed=9364765;  
RA Mannen H., Tsai S.-C.-M., Knapke J.S., Li W.-H., Li S.-L.;  
RT "The cDNA cloning and molecular evolution of reptile and pigeon  
lactate dehydrogenase isozymes."  
RL Mol. Biol. Evol. 14:1081-1087(1997).  
CC -1- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) = pyruvate + NADH.  
CC -1- PATHWAY: Anaerobic glycolysis; final step.  
CC -1- SUBUNIT: Homotrimer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Belongs to the LDH/MDH superfamily. LDH family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.

DR EMBL; L76362; AAD46976.1; -, mRNA.  
DR HSSP; P00339; 9LDT.  
DR SMR; Q9PMO7; 1-331.  
DR InterPro; IPR011304; L-LDH-NAD.  
DR InterPro; IPR001557; L\_LDH\_MDH.  
DR InterPro; IPR001236; Idh.  
DR PANTHER; PTHR11540; Idh; 1.  
DR Pfam; PF02866; Idh\_1\_C; 1.  
DR Pfam; PF00056; Idh\_1\_N; 1.  
DR PRISF; PRSF000102; Lac\_mal\_DH; 1.  
DR PRINTS; PRO0086; LLDHOGNASE.  
DR TIGRFAMs; TIGR01771; L-LDH-NAD; 1.  
DR PROSITE; PS00064; L\_LDH; 1.  
KW Glycolysis; Multigene family; NAD; Oxidoreductase.  
FT INIT MET 0  
FT NP BIND 28 56  
FT ACT SITE 192 192  
FT BINDING 98 98 Proton acceptor (By similarity).  
FT BINDING 105 105 NAD (By similarity).  
FT BINDING 137 137 Substrate (By similarity).  
FT BINDING 168 168 NAD or substrate (By similarity).  
FT BINDING 247 247 Substrate (By similarity).  
SQ SEQUENCE 331 AA; 36396 MW; FZAD59FEFB94CDBA CRC64;

Query Match 88.4%; Score 84; DB 1; Length 331;  
Best Local Similarity 93.3%; Pred. No. 0.00016;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LIHNVHKERHAHAN 16  
| | | | | : | | | |  
Db 6 LIHNVHKERSHAN 20

RESULT 5  
LDHA\_PELSU STANDARD; PRT; 331 AA.  
AC O98SLQ;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE l-lactate dehydrogenase A chain (EC 1.1.1.27) (LDH-A).  
GN Name=LDHA;  
OS Pelodiscus sinensis japonicus (Chinese soft-shelled turtle).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Testudines; Cryptodira; Trionychoidae; Trionychidae; Pelodiscus.  
CX NCBI\_TaxID=34908;  
NN [1]

RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Muscle;  
RX MEDLINE=21579788; PubMed=11722846; DOI=10.1016/S0378-1119(01)00746-6;  
RA Liao C.-H., Ho W.-Z., Huang H.-W., Kuo C.-H., Lee S.-C., Li S.S.-L.;  
RT "Lactate dehydrogenase genes of catman and Chinese soft-shelled  
turtle, with emphasis on the molecular phylogenetics and evolution of  
reptiles.";  
RL Gene 279:63-67(2001).  
CC -1- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) = pyruvate + NADH.  
CC -1- PATHWAY: Anaerobic glycolysis; final step.  
CC -1- SUBUNIT: Homotrimer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Belongs to the LDH/MDH superfamily. LDH family.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR EMBL; AF363794; AAK37572.1; -; mRNA.  
DR HSSP; P00339; 9LDT.  
DR InterPro; IPR011304; L-LDH-NAD.  
DR InterPro; IPR001557; L\_LDH\_MDH.  
DR InterPro; IPR001236; Ldh.  
DR PANTHER; PTHR11540; Ldh; 1.  
DR Pfam; PF02866; Ldh\_1\_C; 1.  
DR Pfam; PF00056; Ldh\_1\_N; 1.  
DR PIRSF; PIRSF00102; Lac\_mal\_DH; 1.  
DR PRINTS; PRO0086; LLDHDEGNASE.  
DR TIGRfam; TIGR01771; L-LDH-NAD; 1.  
DR PROSITE; PS00064; L\_LDH; 1.  
KW Glycolysis; Multigene family; NAD; Oxidoreductase.  
FT INIT MET 0 0  
FT NP BIND 281 56 NAD (By similarity).  
FT ACT\_SITE 192 192 Proton acceptor (By similarity).  
FT BINDING 98 98 NAD (By similarity).  
FT BINDING 105 105 Substrate (By similarity).  
FT BINDING 137 137 NAD or substrate (By similarity).  
FT BINDING 168 168 Substrate (By similarity).  
FT BINDING 247 247 Substrate (By similarity).  
SQ SEQUENCE 331 AA; 36488 MW; 2D9340A5D4F199 CRC64;  
Query Match 80.0%; Score 76; DB 1; Length 331;  
Best Local Similarity 86.7%; Pred. No. 0.003;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 LIHNVKKEHSHAHN 16  
ID |||||:||||  
AC 06SSM1;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE L-lactate dehydrogenase A chain (EC 1.1.1.27) (LDH-A).  
GN Name=LDH-A;  
OS Trachelemys scripta (Red-eared slider turtle) (Pseudemys scripta).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Testudines; Cryptodira; Testudinoidea; Emydidae; Trachemys.  
OX NCB1\_TaxID=34903;  
RN (1)  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Subsp. elegans; TISSUE=Muscle;  
RX MEDLINE=98031373; PubMed=9364765;  
RA Mammen H., Teol S.C.-W., Krushkal J.S., Li W.-H., Li S.S.-L.;  
RT "The cDNA cloning and molecular evolution of reptile and pigeon  
RT lactate dehydrogenase isozymes.";  
RL Mol. Biol. Evol. 14:1081-1087(1997).

CC -1- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) = pyruvate + NADH.  
CC -1- PATHWAY: Anaerobic glycolysis; final step.  
CC -1- SUBUNIT: Homotrimer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Belongs to the LDH/MDH superfamily. LDH family.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR EMBL; U79953; AAD46979.1; -; mRNA.  
DR HSSP; P00339; 9LDT.  
DR InterPro; IPR011304; L-LDH-NAD.  
DR InterPro; IPR001557; L\_LDH\_MDH.  
DR InterPro; IPR001236; Ldh.  
DR PANTHER; PTHR11540; Ldh; 1.  
DR Pfam; PF02866; Ldh\_1\_C; 1.  
DR Pfam; PF00056; Ldh\_1\_N; 1.  
DR PIRSF; PIRSF00102; Lac\_mal\_DH; 1.  
DR PRINTS; PRO0086; LLDHDEGNASE.  
DR TIGRfam; TIGR01771; L-LDH-NAD; 1.  
DR PROSITE; PS00064; L\_LDH; 1.  
KW Glycolysis; Multigene family; NAD; Oxidoreductase.  
FT INIT MET 0 0  
FT NP BIND 28 56 NAD (By similarity).  
FT ACT\_SITE 192 192 Proton acceptor (By similarity).  
FT BINDING 98 98 NAD (By similarity).  
FT BINDING 105 105 Substrate (By similarity).  
FT BINDING 137 137 NAD or substrate (By similarity).  
FT BINDING 168 168 Substrate (By similarity).  
FT BINDING 247 247 Substrate (By similarity).  
SQ SEQUENCE 331 AA; 36510 MW; A7F090A4EE50A00E CRC64;  
Query Match 80.0%; Score 76; DB 1; Length 331;  
Best Local Similarity 86.7%; Pred. No. 0.003;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 LIHNVKKEHSHAHN 16  
ID |||||:||||  
AC 06SSM1;  
DT 05-JUL-2004 (TReMBLrel. 27, Created)  
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
DE L-lactate dehydrogenase A.  
GN Name=LDH-A;  
OS Macroclemys temminckii (Alligator snapping turtle).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Testudines; Cryptodira; Testudinoidea; Chelydridae; Macroclemys.  
OX NCB1\_TaxID=75444;  
RN (1)  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Heart;  
RA Ho W.-Z., Li S.S.-L.;  
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY464556; AAR27956.1; -; mRNA.  
DR HSSP; P6115; 1A5Z.  
DR GO; GO:0005737; C:cytoplasm; IEA.  
DR GO; GO:0004459; F:L-lactate dehydrogenase activity; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0019642; P:anaerobic glycolysis; IEA.  
DR GO; GO:0006100; P:tricarboxylic acid cycle intermediate metab. . .; IEA.  
DR InterPro; IPR011304; L-LDH-NAD.  
DR InterPro; IPR001236; Ldh.  
DR InterPro; IPR001557; L\_LDH\_MDH.  
DR Pfam; PF02866; Ldh\_1\_C; 1.

DR Pfam: PF00056; Ldh\_1\_N; 1.  
 DR PIRSF: PIRSF00102; Lac dehydrog; 1.  
 DR PRINTS: PR00086; LLDHGRNASE.  
 DR TIGRFAMs: TIGR01771; L-LDH-NAD; 1.  
 DR PROSITE: PS00064; L\_LDH; UNKNOWN 1.  
 SQ SEQUENCE 332 AA; 36685 MW; F6CCF7CDBC6D8381 CRC64;

Query Match 1 80.0%; Score 76; DB 2; Length 332;  
 Best Local Similarity 86.7%; Pred. No. 0.003;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LIIHWKEKSHAHN 16  
 |||||  
 Db 7 LIIHWKEKSHAHN 21

## RESULT 8

06SSM4\_9SAUR PRELIMINARY; PRT; 332 AA.

AC 06SSM4;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE L-lactate dehydrogenase A.  
 GN Name=Ldh-A;

OS Carettocheilys insculpta (pitted-shelled turtle).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Testudines; Cryptodira; Trionychidae; Carettocheilys;  
 OC Carettocheilys.

OX NCBI\_TaxId=44489;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Muscle;  
 RA Ho W.-Z., Li S.S.-L.;  
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY464553; AAR27953.1; -; mRNA.

DR HSSP: P16115; 1A5Z.  
 DR GO: GO:0005737; Cytoplasm; IEA.  
 DR GO: GO:0004459; F:L-lactate dehydrogenase activity; IEA.

DR GO: GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO: GO:0019642; P:anaerobic glycolysis; IEA.  
 DR GO: GO:0006100; P:tricarboxylic acid cycle intermediate metab. . .; IEA.

DR InterPro: IPR011304; L-LDH-NAD.  
 DR InterPro: IPR01236; Ldh.  
 DR InterPro: IPR01557; L\_LDH\_MDH.

DR Pfam: PF02866; Ldh\_1\_C; 1.  
 DR Pfam: PF00056; Ldh\_1\_N; 1.  
 DR PIRSF: PIRSF00102; Lac dehydrog; 1.

DR PRINTS: PR00086; LLDHGRNASE.  
 DR TIGRFAMs: TIGR01771; L-LDH-NAD; 1.  
 DR PROSITE: PS00064; L\_LDH; UNKNOWN 1.  
 SQ SEQUENCE 332 AA; 36591 MW; 84622AB841F97A4B CRC64;

Query Match 80.0%; Score 76; DB 2; Length 332;  
 Best Local Similarity 86.7%; Pred. No. 0.003;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LIIHWKEKSHAHN 16  
 |||||  
 Db 7 LIIHWKEKSHAHN 21

RESULT 9  
 06SSM6\_APAFE PRELIMINARY; PRT; 332 AA.

AC 06SSM6;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE L-lactate dehydrogenase A.  
 GN Name=Ldh-A;  
 OS Apalone ferox (Florida softshell turtle).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Testudines; Cryptodira; Trionychidae; Trionychidae; Apalone.  
 OX NCBI\_TaxId=100213;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Muscle;  
 RA Ho W.-Z., Li S.S.-L.;  
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY464551; AAR27951.1; -; mRNA.  
 DR HSSP: P16115; 1A5Z.  
 DR GO: GO:0005737; Cytoplasm; IEA.

DR GO: GO:0004459; F:L-lactate dehydrogenase activity; IEA.  
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.

DR GO: GO:0019642; P:anaerobic glycolysis; IEA.  
 DR GO: GO:0006100; P:tricarboxylic acid cycle intermediate metab. . .; IEA.

DR InterPro: IPR011304; L-LDH-NAD.  
 DR InterPro: IPR01236; Ldh.  
 DR InterPro: IPR01557; L\_LDH\_MDH.

DR Pfam: PF02866; Ldh\_1\_C; 1.  
 DR Pfam: PF00056; Ldh\_1\_N; 1.  
 DR PIRSF: PIRSF00102; Lac dehydrog; 1.

DR PRINTS: PR00086; LLDHGRNASE.  
 DR TIGRFAMs: TIGR01771; L-LDH-NAD; 1.  
 DR PROSITE: PS00064; L\_LDH; UNKNOWN 1.  
 SQ SEQUENCE 332 AA; 36611 MW; 29BDDCE1C70AF1E CRC64;

Query Match 80.0%; Score 76; DB 2; Length 332;  
 Best Local Similarity 86.7%; Pred. No. 0.003;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LIIHWKEKSHAHN 16  
 |||||  
 Db 7 LIIHWKEKSHAHN 21

## RESULT 10

06SSM2\_9SAUR PRELIMINARY; PRT; 332 AA.

AC 06SSM2;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE L-lactate dehydrogenase A.  
 GN Name=Ldh-A;

OS Chelodina siebenrocki.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Testudines; Pleurodira; Chelididae; Chelodina.

OX NCBI\_TaxId=257467;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.

RA Ho W.-Z., Li S.S.-L.;  
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY464555; AAR27955.1; -; mRNA.

DR HSSP: P16115; 1A5Z.  
 DR GO: GO:0005737; Cytoplasm; IEA.  
 DR GO: GO:0004459; F:L-lactate dehydrogenase activity; IEA.

DR GO: GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO: GO:0019642; P:anaerobic glycolysis; IEA.  
 DR GO: GO:0006100; P:tricarboxylic acid cycle intermediate metab. . .; IEA.

DR InterPro: IPR011304; L-LDH-NAD.  
 DR InterPro: IPR01236; Ldh.  
 DR InterPro: IPR01557; L\_LDH\_MDH.

DR Pfam: PF02866; Ldh\_1\_C; 1.  
 DR Pfam: PF00056; Ldh\_1\_N; 1.  
 DR PIRSF: PIRSF00102; Lac dehydrog; 1.

DR PRINTS: PR00086; LLDHGRNASE.  
 DR TIGRFAMs: TIGR01771; L-LDH-NAD; 1.  
 DR PROSITE: PS00064; L\_LDH; UNKNOWN 1.  
 SQ SEQUENCE 332 AA; 36612 MW; C3F619A40876EB0B CRC64;

Query Match 76.8%; Score 73; DB 2; Length 332;  
 Best Local Similarity 80.0%; Pred. No. 0.0084;  
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

Oy      2 LIHNVKEHEHAH 16
      |||||
Db      7 LIQNVKHEHSHSN 21

RESULT 11
O6YL22 IGUG PRELIMINARY; PRT; 332 AA.
AC O6YL22;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Lactate dehydrogenase A.
OC Iguana iguana (Common iguana).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidotesauria; Squamata; Iguania; Iguanidae; Iguaninae; Iguana.
OC NCBI_TaxID=8517;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hau C.-H., Li S.S.-L.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AI130249; AAN05098.1; -, mRNA.
DR HSSP; P16115; IASZ.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0004459; F:L-lactate dehydrogenase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0019642; P:anaerobic glycolysis; IEA.
DR GO; GO:0006100; P:tricarboxylic acid cycle intermediate metab. .; IEA.
DR InterPro; IPR011304; L-LDH-NAD.
DR InterPro; IPR001356; Ldh.
DR Pfam; PF02866; Ldh_1_N; 1.
DR PIRSF; PIRSF00102; Lac dehydrog; 1.
DR PRINTS; PR00086; LLDHNRGNASE.
DR TIGRFAMs; TIGR01771; L-LDH-NAD; 1.
DR PROSITE; PS00064; L_LDH; UNKNOWN 1.
SQ SEQUENCE 332 AA; 36532 MW; 70E9FA56DFA384D3 CRC64;

Query Match 75.8%; Score 72; DB 2; Length 332;
Best Local Similarity 86.7%; Pred. No. 0.012;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy      2 LIHNVKEHEHAH 16
      |||||
Db      7 LIQNVKHEHSHSN 21

RESULT 12
LDHA PYTRG STANDARD; PRT; 331 AA.
AC O9W1D3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE L-lactate dehydrogenase A chain (EC 1.1.1.27) (LDH-A).
GN Name=LDHA;
OS Python regius (Ball python) (Royal python).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidotesauria; Squamata; Scleroglossa; Serpentes; Hemophidia;
OC Pythonidae; Python.
OC NCBI_TaxID=51751;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA TISSUE=Muscle;
MEDLINE=99439677; PubMed=10508547; DOI=10.1006/mpcv.1999.0640;
Mamen H., Li S.S.-L.;
"molecular evidence for a clade of turtles.";
RL Mol. Phylogenet. Evol. 13:144-148(1999).
CC -1- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) = pyruvate + NADH.
CC -1- PATHWAY: Anaerobic glycolysis; final step.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
```

```

CC -1- SIMILARITY: Belongs to the LDH/MDH superfamily. LDH family.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL; AF072585; AAD41642.1; -, mRNA.
DR HSSP; P00338; 1110.
DR InterPro; IPR011304; L-LDH-NAD.
DR InterPro; IPR001557; L_LDH_MDH.
DR InterPro; IPR001236; Ldh.
DR PANTHER; PTHR11540; Ldh; 1.
DR Pfam; PF02866; Ldh_1_C; 1.
DR Pfam; PF00056; Ldh_1_N; 1.
DR PIRSF; PIRSF000102; Lac_mal_DH; 1.
DR PRINTS; PR00086; LLDHNRGNASE.
DR TIGRFAMs; TIGR01771; L-LDH-NAD; 1.
DR PROSITE; PS00064; L_LDH; 1.
KM Glycolysis; Multigene family; NAD; Oxidoreductase.
FT INIT_MET 0
FT NP_BIND 28
FT ACT_SITE 192 192
FT BINDING 98 98
FT BINDING 105 105
FT BINDING 137 137
FT BINDING 168 168
FT BINDING 247 247
SQ SEQUENCE 331 AA; 36573 MW; 4DAAF72F88A0B481 CRC64;

Query Match 69.5%; Score 66; DB 1; Length 331;
Best Local Similarity 80.0%; Pred. No. 0.096;
Matches 12; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy      2 LIHNVKEHEHAH 16
      |||||
Db      6 LIQNVKHEHSHSN 20

RESULT 13
LDHA SCENO STANDARD; PRT; 331 AA.
AC P79912;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE L-lactate dehydrogenase A chain (EC 1.1.1.27) (LDH-A).
GN Name=LDHA;
OS Sceloporus woodi (Florida scrub lizard).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidotesauria; Squamata; Iguania; Iguanidae; Phrynosomatinae;
OC Sceloporus.
OC NCBI_TaxID=59726;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96257240; PubMed=8666293; DOI=10.1016/0378-1119(96)00180-1;
Mamen H., Teot S.C.-M., Pickford D.B., Donald J.A., Guillelte L.J.,
Li S.S.-L.;
"Sequences of the lizard cDNAs encoding lactate dehydrogenase (LDH)
RT isozymes A (muscle) and B (heart).";
RL Gene 171:303-304(1996).
CC -1- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) = pyruvate + NADH.
CC -1- PATHWAY: Anaerobic glycolysis; final step.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the LDH/MDH superfamily. LDH family.
CC -1- CAUTION: Was originally (Ref.1) thought to originate from
CC S.undulatus.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```

CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL, U78410; AAB53025.1; -, mRNA.
DR HSSP, P00338; 1110.
DR SMR, P79912; 1-331.
DR InterPro, IPR011304; L-LDH-NAD.
DR InterPro, IPR001557; L-LDH_MDH.
DR InterPro, IPR001236; 1dh.
DR PANTHER, PTHR11540; 1dh; 1.
DR Pfam, PF02866; Ldh_1_C; 1.
DR Pfam, PF00056; Ldh_1_N; 1.
DR PIRSF, PIRSF00102; Lac mal DH; 1.
DR PRINTS, PRO0086; LLDHGRNASE.
DR TIGRPFAM, TIGR01771; L-LDH-NAD; 1.
DR PROSITE, PS00064; L_LDH; 1.
KW Glycolysis; Multigene family; NAD; Oxidoreductase.
FT INIT MET 0 0 By similarity.
FT NP BIND 28 56 NAD (By similarity).
FT ACT SITE 192 192 Proton acceptor (By similarity).
FT BINDING 98 98 NAD (By similarity).
FT BINDING 105 105 Substrate (By similarity).
FT BINDING 137 137 NAD or substrate (By similarity).
FT BINDING 168 168 Substrate (By similarity).
FT BINDING 247 247 Substrate (By similarity).
SQ SEQUENCE 331 AA; 36411 MW; E197020968123A23 CRC64;

Query Match 66.3%; Score 63; DB 1; Length 331;
Best Local Similarity 80.0%; Pred. No. 0.27;
Matches 12; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 L1HNHKEEHAHNN 16
DB 6 L1VNVKEEOPHANN 20

RESULT 14
LDHA_SCEUN STANDARD; PRT; 331 AA.
ID LDHA_SCEUN STANDARD; PRT; 331 AA.
AC Q9W7L5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE L-lactate dehydrogenase A chain (EC 1.1.1.27) (LDH-A).
GN Name=LDHA;
OS Sceloporus undulatus (Eastern fence lizard) (Skink).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Iguanidae; Phrynosomatinae;
OC Sceloporus;
OC NCBI_TaxID=8520;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99439677; PubMed=10508547; DOI=10.1006/mpev.1999.0640;
RA Mannen H., Li S.S.-L.;
RL "Molecular evidence for a clade of turtles.";
Mol. Phylogene. Evol. 13:144-148(1999).
CC -1- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) = pyruvate + NADH.
CC -1- PATWAY: Anaerobic glycolysis; final step.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the LDH/MDH superfamily. LDH family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL, AF072583; AAD41640.1; -, mRNA.
DR HSSP, P00339; 91DT.
DR InterPro, IPR011304; L-LDH-NAD.
DR InterPro, IPR001557; L_LDH_MDH.

```

```

DR InterPro, IPR001236; 1dh.
DR PANTHER, PTHR11540; 1dh; 1.
DR Pfam, PF02866; Ldh_1_C; 1.
DR Pfam, PF00056; Ldh_1_N; 1.
DR PIRSF, PIRSF00102; Lac mal DH; 1.
DR PRINTS, PRO0086; LLDHGRNASE.
DR TIGRPFAM, TIGR01771; L-LDH-NAD; 1.
DR PROSITE, PS00064; L_LDH; 1.
KW Glycolysis; Multigene family; NAD; Oxidoreductase.
FT INIT MET 0 0 By similarity.
FT NP BIND 28 56 NAD (By similarity).
FT ACT SITE 192 192 Proton acceptor (By similarity).
FT BINDING 98 98 NAD (By similarity).
FT BINDING 105 105 Substrate (By similarity).
FT BINDING 137 137 NAD or substrate (By similarity).
FT BINDING 168 168 Substrate (By similarity).
FT BINDING 247 247 Substrate (By similarity).
SQ SEQUENCE 331 AA; 36496 MW; A69A0ADEDCAPZ68 CRC64;

Query Match 65.3%; Score 62; DB 1; Length 331;
Best Local Similarity 73.3%; Pred. No. 0.39;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 L1HNHKEEHAHNN 16
DB 6 L1EKVTEEHPHANN 20

RESULT 15
O91TX1_ARATH PRELIMINARY; PRT; 662 AA.
ID O91TX1_ARATH PRELIMINARY; PRT; 662 AA.
AC O91TX1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similarity to poly(A)-binding protein.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20277480; PubMed=10819329;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:131-135(2000).
RN (2)
RP NUCLEOTIDE SEQUENCE.
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL, AB024033; BAB02407.1; -, Genomic DNA.
DR EMBL, AF002044; BAB02407.1; JOINED; Genomic DNA.
DR HSSP, P33240; 1PT.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0000398; F:nuclear RNA splicing, via spliceosome; IEA.
DR InterPro, IPR012677; a b plat nuc_bd.
DR InterPro, IPR00504; RNPL_RNA_Bd.
DR SMART, SM0076; RRM_1; 1.
DR SMART, SM00360; RRM; 1.
DR PROSITE, PS50102; RRM; 1.
SQ SEQUENCE 662 AA; 73082 MW; 04941480A95560C CRC64;

Query Match 62.1%; Score 59; DB 2; Length 662;
Best Local Similarity 56.2%; Pred. No. 2.4;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 H1HNHKEEHAHNN 16
DB 157 H1HNHKEEHAHNN 172

```

```

RESULT 16
09C7B4_ARATH PRELIMINARY; PRT; 674 AA.
ID 09C7B4;
AC 09C7B4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein T2E22.5.
GN Name: T2E22.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Malt R., Roming C.M., Koo H., Fujii C.Y., Ullrich T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nieman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC069474; AAG51005.1; -; Genomic_DNA.
DR HSSP; P33240; IPT.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0000398; F:nuclear mRNA splicing; IEA.
DR InterPro; IPR012677; a b plat nuc_bd.
DR InterPro; IPR005054; RNP_RNA_bd.
DR Pfam; PF00076; RRM_1; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS0102; RRM; 1.
KM Hypothetical protein.
SQ
SEQUENCE 674 AA; 7446 MW; 2D6A1FB4BC98D75D CRC64;

Query Match 62.1%; Score 59; DB 2; Length 674;
Best Local Similarity 56.2%; Pred. No. 2.4;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 HLHNVHKEBHAAHN 16
DB 157 HKIHNYEKDHKRRHN 172

RESULT 17
05E729_VIBF1 PRELIMINARY; PRT; 163 AA.
ID 05E729;
AC 05E729;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Urease accessory protein UreB.
GN OrderedAccession=VF0672;
OS Vibrio fischeri (strain ATCC 700601 / ES114).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=312309;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA PubMed=15703294; DOI=10.1073/pnas.0409900102;
RA Ruby E.G., Urbanowski M., Campbell J., Dunn A., Faini M., Gunsalus R.,
RA Loeferich P., Lupp C., McCann J., Millikan D., Schaefer A., Stabb E.,
RA Stevens A., Vialick K., Whistler C., Greenberg E.P.;
RT "Complete genome sequence of Vibrio fischeri: a symbiotic bacterium
with pathogenic congeners."
RL Proc. Natl. Acad. Sci. U.S.A. 102:3004-3009 (2005).
DR EMBL; CP000020; AAM85167.1; -; Genomic_DNA.
DR InterPro; IPR007864; UreG_C.
DR InterPro; IPR004029; UreG_N.
DR Pfam; PF05194; UreG_C_1.
DR Pfam; PF02814; UreG_N_1.
KM Complete proteome.
SQ
SEQUENCE 163 AA; 18287 MW; E674232BB37AD402 CRC64;

```

```

Query Match 57.4%; Score 54.5; DB 2; Length 163;
Best Local Similarity 52.6%; Pred. No. 2.4;
Matches 10; Conservative 2; Mismatches 4; Indels 3; Gaps 1;

QY 1 HLHNVHKEBHAAHN 16
DB 6 HLHVVHHHDEHHNEBHAAHN 24

RESULT 18
07SCY2_NEUCR PRELIMINARY; PRT; 1032 AA.
ID 07SCY2;
AC 07SCY2;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=NCU02752.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Puccelli S., Rehman B.,
RA Ekins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Mashburne M.,
RA Seltzer M., Kline J.A., Braun E.L., Zelter A., Schulte D.,
RA Kolte G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal T., Kamysheva M., Maccell E., Blake C., Rudd S., Fishman D.,
RA Kiyotoku S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmari S.A.,
RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Selter S., Dunlap J., Radford A., Aramayo R.,
RA Nardig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.,
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
RL Nature 0:0-0 (2003).
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL; AABX0100073; BAA34609.1; -; Genomic_DNA.
DR HSSP; P12351; IHWT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; P:metal ion binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006355; P:regulation of transcription; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR007219; Fungal trans.
DR InterPro; IPR01138; Fungal Transcrp_N.
DR Pfam; PF04082; Fungal trans; 1.
DR Pfam; PF00172; Zn_c1ug; 1.
DR PRINTS; PR00054; FUNGALZNCYS.
DR PROSITE; PS00463; ZN2_Cy6_FUNGAL_1; 1.
DR PROSITE; PS00463; ZN2_Cy6_FUNGAL_2; 1.
KM DNA-binding; Hypothetical protein; Metal-binding; Nuclear protein;
KM Transcription; Transcription regulation; Zinc.
SQ
SEQUENCE 1032 AA; 112225 MW; 395D61F55B33AB8 CRC64;

Query Match 56.8%; Score 54; DB 2; Length 1032;
Best Local Similarity 43.8%; Pred. No. 22;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 HLHNVHKEBHAAHN 16
DB 798 HLCHQHCHQHCHQHCH 813

```



```

RESULT 19
Q8T990.DROME
ID Q8T990.DROME PRELIMINARY; PRT; 725 AA.
AC Q8T990;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE S001032P.
GN ORFNames=CG4136;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN
RP NUCLEOTIDE SEQUENCE.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AY070708; AL448179.1; -, mRNA.
DR HSSP; P06071; 1FJL.
DR FLYBase; FBgn0029775; CG4136.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006555; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR007104; Paired_homeo.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PRO0024; HOMEBOX.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1, 1.
DR PROSITE; PS00071; HOMEBOX_2, 1.
DR DNA-binding; Developmental protein; Homeobox; Nuclear protein;
KW Transcription; Transcription regulation.
SQ SEQUENCE 725 AA; 75302 MW; A1D40CA80825E40F CRC64;

Query Match 55.8%; Score 53; DB 2; Length 725;
Best Local Similarity 53.3%; Pired. No. 21;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 HLHNYHKEEHAHAH 15
Db 206 HHAHHAAQAHAAH 220

RESULT 20
Q9W4B2.DROME
ID Q9W4B2.DROME PRELIMINARY; PRT; 837 AA.
AC Q9W4B2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE CG4136-PA.
GN Name=CG4136; ORFNames=CG4136;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,

```

```

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Adril J.F., Agbayani A., An H.-O., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintsov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murty D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Murty D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RL melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminiker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RL a genomics perspective."
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminiker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Beutenkourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RL systematic review."
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN
RP NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;

```

```
RT "Drosophila melanogaster release 4 sequence."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RG Flybase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AE003434; AAF4604.1; -; Genomic_DNA.
DR HSSP; P06601; 1FVL.
DR Flybase; FBgn0029775; CG4136.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription; IEA.
DR InterPro; IPR012287; Homeobox.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR007104; Paired_homeo.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR Drosophila; Developmental protein; Homeobox; Nuclear protein; Transcription; Transcription regulation.
KM Drosophila; Developmental protein; Homeobox; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 837 AA; 8686 MW; 54955738CD7203D9 CRC64;

Query Match 55.8%; Score 53; DB 2; Length 837;
Best Local Similarity 53.3%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 HLIHNVKEBHAAH 15
DB 318 HHAHHAAQAHAAH 332

RESULT 21
Q5UID_DROME PRELIMINARY; PRT; 837 AA.
AC Q5UID;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE SD0647D.
GN Name=CG4136;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OC NCB1_TaxId=7227;
RN NCB1_TaxId=7227;
RP NUCLEOTIDE SEQUENCE.
RA Stapleton M., Carlson J., Chavez C., Friese B., George R., Pacleb J., Park S., Wan K., Yu C., Rubin G.M., Calniker S.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; BT015953; AAV36838.1; -; mRNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription; DNA-dependent; IEA.
DR InterPro; IPR01356; Homeobox.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR007104; Paired_homeo.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KM DNA-binding; Developmental protein; Homeobox; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 837 AA; 8685 MW; 51A6A73D3374F27 CRC64;

Query Match 55.8%; Score 53; DB 2; Length 837;
Best Local Similarity 53.3%; Pred. No. 25;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 HLIHNVKEBHAAH 15
DB 318 HHAHHAAQAHAAH 332

RESULT 22
Q6C9B8_YARLI PRELIMINARY; PRT; 717 AA.
AC Q6C9B8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similarity
GN OrderedCusNames=YAL10D12430g;
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Dipodascaceae; Yarrowia.
OC NCB1_TaxId=4952;
RN NCB1_TaxId=4952;
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Duret P., Casaregola S., Talla E., Lafontaine I., de Montigny J., Marcq C., Neuvéglise C., Talla B., Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V., Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C., Boirame A., Boyer J., Catolico L., Contamoli F., de Darvar A., Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Gropfi A., Hantreya F., Hennequin C., Jauniaux N., Joyet P., Kachouri R., Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Miller H., Nicoud J.-M., Nikolski M., Ozias S., Ozier-Kalogeropoulos O., Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A., Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B., Zenou-Meyer M., Zivanovic Y., Bolestin-Fukuhara M., Thierry A., RA Bouchier C., Caudron B., Scaupell C., Galliard C., Weissenbach J., RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
DR EMBL; CR382130; CAG80932.1; -; Genomic_DNA.
DR GO; GO:0005676; F:nucleic acid binding; IEA.
DR InterPro; IPR01374; R3H.
DR Pfam; PF01424; R3H; 1.
KM Complete proteome.
SQ SEQUENCE 717 AA; 76303 MW; 9020C83728E15756 CRC64;

Query Match 54.7%; Score 52; DB 2; Length 717;
Best Local Similarity 53.3%; Pred. No. 30;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 HLIHNVKEBHAAH 15
DB 643 HMPHQAEMGHAAH 657

RESULT 23
Q7S3X5_NEUCR PRELIMINARY; PRT; 1018 AA.
AC Q7S3X5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=NCU0214.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OC NCB1_TaxId=5141;
RN NCB1_TaxId=5141;
RP NUCLEOTIDE SEQUENCE.
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D., Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
```

RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
RA Qui D., Iankiev P., Pedersen D., Nelson M., Washburne M.,  
RA Seltenreich C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,  
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,  
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,  
RA Kamal M., Kamysseil M., Mauceli E., Bielek C., Rudd S., Friesman D.,  
RA Kryzofowa S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,  
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,  
RA Desouza C.C., Glaes L., Orbach M.J., Berglund J., Voelker R.,  
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,  
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,  
RA Paulsen I., Sachs M.S., Lander E.S., Nisbaum C., Birren B.,  
RT "The genome sequence of the Filamentous Fungus Neurospora crassa.",  
RL Nature 0:0-0(2003).  
CC -!- CAUTION: The sequence shown here is derived from an  
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
preliminary data.  
CC EMBL: AABX01000392; EMBL: AABX01000392; EMBL: AABX01000392;  
DR EMBL: AABX01000392; EMBL: AABX01000392; EMBL: AABX01000392;  
DR GO: GO:0005634; C:nucleus; IEA.  
DR GO: GO:0046872; F:metal ion binding; IEA.  
DR GO: GO:0003700; F:transcription factor activity; IEA.  
DR GO: GO:0008270; F:zinc ion binding; IEA.  
DR GO: GO:0006355; P:regulation of transcription; DNA-dependent; IEA.  
DR GO: GO:0006350; P:transcription; IEA.  
DR InterPro: IPR007219; Fungal\_trans.  
DR InterPro: IPR001138; Fungal\_trans.  
DR Pfam: PF04082; Fungal\_trans; 1.  
DR Pfam: PF00172; Zn\_c1ub; 1.  
DR PRINTS: PR00054; FUNGAL\_ZINC.  
DR PROSITE: PS00463; ZN2\_Cy6\_FUNGAL\_1; 1.  
DR PROSITE: PS00463; ZN2\_Cy6\_FUNGAL\_2; 1.  
KM DNA-binding: Hypothetical protein; Metal-binding; Nuclear protein;  
KW Transcription; Transcriptional regulation; Zinc.  
SQ SEQUENCE 1018 AA; 110551 MW; 9774FC2A86930B92 CRC64;  
  
Query Match 54.7%; Score 52; DB 2; Length 1018;  
Best Local Similarity 53.3%; P: 44;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 HLINHYKEHNAH 15  
Db 908 YLLHHHHHHNAH 922  
  
RESULT 24  
Q8INR1 DROME PRELIMINARY; PRT; 339 AA.  
ID Q8INR1 DROME PRELIMINARY; PRT; 339 AA.  
AC Q8INR1;  
DT 01-MAR-2003 (TEMBLrel. 23, Created)  
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)  
DE CG9610-PB, Isoform B.  
GN Name=Pxmx; ORFNames=CG9610;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
NCBI TaxID=7227;  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RP NUCLEOTIDE SEQUENCE.  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer V., Champe M., Pfeiffer B.D.,  
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glorick A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Iasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Izu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Morklov G., Mshelishvili N.V., Moberg C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
RA Palazolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverly T., Murry D.M., Nelson C.R.,  
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila  
melanogaster euchromatic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.B., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the Drosophila melanogaster euchromatin:  
a genomic perspective.";  
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Mista S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kinkler J.S., Milburn G.H., Prochick S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
RA Betencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter U., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
systematic review.";  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RA Berkeley Drosophila Genome Project;  
RG Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
RA Hoskins R., Stapleton M., Pacle J., Park S., Svirskas R., Smith E.,  
RA Yu C., Rubin G.;  
RT "Drosophila melanogaster release 4 sequence.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP NUCLEOTIDE SEQUENCE.  
RA Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).



```
RESULT 26
Q8IDA6_PLA67
ID Q8IDA6_PLA67 PRELIMINARY; PRT; 2126 AA.
AC Q8IDA6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein MAL13P1.298.
DE Name=MAL13P1.298;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryotes; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3D7;
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Barrigan M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrett B.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844509; CAD52719.1; -; Genomic_DNA.
KM Hypothetical protein.
SQ SEQUENCE 2126 AA; 249758 MW; 11968B36CF207FF3 CRC64;

Query Match 53.7%; Score 51; DB 2; Length 2126;
Best Local Similarity 57.1%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 HNVHKEHHAAH 16
DB 1304 HNHHEHHEHHEH 1317

RESULT 27
Q9FUB1_ARATH
ID Q9FUB1_ARATH PRELIMINARY; PRT; 245 AA.
AC Q9FUB1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE PRL1-interacting factor 1 (Fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsie.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98438452; PubMed=9765207;
RA Nemeth K., Salchert K., Putnok P., Bhalaria R., Koncz-Kalman Z.,
RA Stankovic-Stangeland B., Bako L., Mathur J., Oekresz L., Stabel S.,
RA Geigenberger P., Slitt M., Redei G.P., Schell J., Koncz C.;
RT "Pleiotropic control of glucose and hormone responses by PRL1, a
RT nuclear WD protein, in Arabidopsis.";
RL Genes Dev. 12:3059-3073(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Salchert K., Koncz C.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF315736; AAC31652.1; -; mRNA.
DR HSSP; P24203; INIT.
DR InterPro; IPR003495; CoBw_.
DR InterPro; IPR011629; CoBw_C.
DR Pfam; PF02492; CoBw_1.
DR Pfam; PF07683; CoBw_C_1.
FT NON_TER 1 1
SQ SEQUENCE 245 AA; 28360 MW; 3A0F7D62A4D118FF CRC64;

Query Match 52.6%; Score 50; DB 2; Length 245;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 HLIHVHKEHHAAH 16
DB 124 HDHNEHEHEHEH 139

RESULT 28
Q6G7G4_STAAS
ID Q6G7G4_STAAS PRELIMINARY; PRT; 326 AA.
AC Q6G7G4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Zinc resistance protein.
DE OrderedLocNames=SAS2049;
GN OrderedLocNames=SAS2049;
OS Staphylococcus aureus (strain MSSA476).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=1521324; DOI=10.1073/pnas.0402521101;
RA Holden M.T.G., Fell E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.B., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Felwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jags J.K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitsch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrett B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; BX571857; CAG43857.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008324; P:cation transporter activity; IEA.
DR GO; GO:0006812; P:cation transport; IEA.
DR InterPro; IPR002524; Cation_efflux.
DR Pfam; PF01545; Cation_efflux; 1.
DR TIGRPFAM; TIGR01297; CDF; 1.
KM Complete proteome.
SQ SEQUENCE 326 AA; 36238 MW; 4AD6A44CCDA209F7 CRC64;

Query Match 52.6%; Score 50; DB 2; Length 326;
Best Local Similarity 66.7%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 HNVHKEHHAAH 15
DB 315 HNHHAHHAAH 326

RESULT 29
Q8NVF2_STAAM
ID Q8NVF2_STAAM PRELIMINARY; PRT; 326 AA.
AC Q8NVF2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CytB protein.
GN Name=cytB; OrderedLocNames=MW2070;
OC Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwano N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; BA000033; BAB95935.1; -; Genomic_DNA.
```

DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0008324; F:cation transporter activity; IEA.  
DR GO; GO:0006812; P:cation transport; IEA.  
DR InterPro: IPR002524; Cation\_efflux.  
DR Pfam: PF01545; Cation\_efflux; 1.  
DR TIGRfam: TIGR01297; CDF; 1.  
DR Complete proteome.  
SQ SEQUENCE 326 AA; 36238 MW; 4AD6A44CDA209F7 CRC64;

Query Match 52.6%; Score 50; DB 2; Length 326;  
Best Local Similarity 66.7%; Pred. No. 25;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 HNVKKEHAAH 15  
Db 315 HNHHAHHAAH 326

RESULT 30  
HRPX\_PLALO STANDARD; PRT; 351 AA.  
ID HRPX\_PLALO  
AC P04929;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE Histidine-rich glycoprotein precursor.  
OS Plasmodium lophurae.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5853;  
RN [1]  
RX NUCLEOTIDE SEQUENCE.  
RA MEDLINE=85061618; PubMed=6095114;  
RA Ravech J.V., Feder R., Pavlovic A., Blobel G.;  
RA "Primary structure and genomic organization of the histidine-rich  
RT protein of the malaria parasite Plasmodium lophurae.";  
RL Nature 312:616-620(1984).  
CC -1- MISCELLANEOUS: In the intraerythrocytic stages of development of  
P. lophurae in ducks, there is a synthesis of a major protein that  
CC accumulates to comprise at least 50% of the cellular mass: the  
CC histidine rich protein.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.

CC EMBL; X01469; CAA25698.1; -; Genomic\_DNA.  
DR PIR; A22692; KGZOH.  
DR HSSP; P13231; IHCR.  
KM Glycoprotein; Malaria; Repeat; Signal.  
FT SIGNAL 1 23  
FT PROPEP 24 47  
FT CHAIN 48 351 Histidine-rich glycoprotein.  
FT REPEAT 59 74 16-1.  
FT REPEAT 75 90 16-2.  
FT REPEAT 91 107 17-1.  
FT REPEAT 108 123 17-2.  
FT REPEAT 124 138 15-1.  
FT REPEAT 139 153 15-2.  
FT REGION 59 90 2 x 16 AA tandem repeats.  
FT REGION 91 123 2 x 17 AA tandem repeats.  
FT REGION 124 153 2 x 15 AA tandem repeats.  
FT REGION 173 351 18 x 10 AA tandem repeats.  
FT CARBOHYD 40 40 N-linked (GlcNAc...) (Probable).  
SQ SEQUENCE 351 AA; 44031 MW; D19A4BD47D890453 CRC64;

Query Match 52.6%; Score 50; DB 1; Length 351;  
Best Local Similarity 50.0%; Pred. No. 27;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 HHHVKEHAAH 16  
| | | | | | | |

Db 123 HAAHHHHHHHHHH 138

RESULT 31  
Q9LMR1\_ARATH PRELIMINARY; PRT; 448 AA.  
ID Q9LMR1;  
AC Q9LMR1;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)  
DE F7H2.7 protein (Atg15730/F7H2.7) (Putative PRL-Interacting factor  
DE l.).  
GN Name=F7H2.7; Synonyms=Atg15730;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosid II; Brassicales; Brassicaceae; Arabidopsie.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Liu S.X., Sakano H., Yu G., Elgu P., Lee J., Lenz C., Pham P.,  
RA Toriumi M., Chan A., Chung M., Goldsmith A., Liu A., Smith A.,  
RA Vayenberg M., Altali H., Brooks S., Buehler E., Chao Q., Conn L.,  
RA Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,  
RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,  
RA Ecker J.R., Federspiel N.A., Theologis A.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,  
RA Bower L., Carninci P., Chang B., Dale J.M., Goldsmith A.D.,  
RA Hayashizaki Y., Ishida J., Jones T., Kamya A., Karlin-Neumann G.,  
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
RA Nguyen M., Onders C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,  
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,  
RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,  
RA Ecker J.R.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,  
RA Bower L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,  
RA Ishida J., Jones T., Kamya A., Karlin-Neumann G., Kawai J.,  
RA Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,  
RA Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,  
RA Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinzaki K.,  
RA Davis R.W., Theologis A., Ecker J.R.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
RN [4]

RP NUCLEOTIDE SEQUENCE.  
RA Yamada K., Chen M.M., Chang C.H., Dale J.M., Huan V.W., Lee J.M.,  
RA Quach H.L., Tang C., Toriumi M., Wallender B.K., Wong C., Wu H.C.,  
RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,  
RA Ishida J., Jones T., Kamya A., Karlin-Neumann G., Kawai J.,  
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,  
RA Southwick A., Tripp M.G., Wu T., Shinzaki K., Davis R.W., Ecker J.R.,  
RA Theologis A.;  
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC034256; AAF82143.1; -; Genomic\_DNA.  
DR EMBL; AY093965; AAM16226.1; -; mRNA.  
DR EMBL; BT000764; AAN31903.1; -; mRNA.  
DR EMBL; AF367261; AAK56250.1; -; mRNA.  
DR PIR; E86291; E86291.  
DR HSSP; P24203; INTI.  
DR InterPro: IPR003495; COB\_M.  
DR InterPro: IPR011629; COB\_M.  
DR Pfam; PF02492; COB\_M; 1.  
DR Pfam; PF07683; COB\_M; 1.  
SQ SEQUENCE 448 AA; 50258 MW; 86CE935F2A069642 CRC64;

Query Match 52.6%; Score 50; DB 2; Length 448;  
Best Local Similarity 50.0%; Pred. No. 35;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;



OY 1 HLHNVHKEEHAHN 16  
Db 327 HDHNEHEHEHEHEH 342

RESULT 32  
OQNE57\_CAEEL PRELIMINARY; PRT; 735 AA.  
ID OQNE57\_CAEEL  
AC OQNE57;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
GN Hypothetical protein Y3986A.1.  
OS ORFNames=Y3986A.1;  
OC Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peleoderinae; Caenorhabditis.  
OX NCBI\_TaxId=6239;  
RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=Bristol N2;  
RG MEDLINE=99069613; PubMed=9851916;  
RT "The C. elegans sequencing consortium;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology";  
RT Science 282:2012-2018(1998).  
RL -1- INTERACTION:  
CC P91419:abu-rl; NDExp=1; InAct=EBI-312701, EBI-313979;  
CC Q18238:cdt27a2.5; NDExp=1; InAct=EBI-312701, EBI-316112;  
CC Q9XVA3:P0862.5; NDExp=1; InAct=EBI-312701, EBI-319064;  
CC O9N552:mes-4; NDExp=1; InAct=EBI-312701, EBI-330178;  
CC O16500:ipm-5; NDExp=1; InAct=EBI-312701, EBI-313836;  
CC Q42406:ipm-5.4; NDExp=1; InAct=EBI-312701, EBI-325936;  
CC Q22174:IP041.2; NDExp=1; InAct=EBI-312701, EBI-311938;  
CC Q45764:R0656.10; NDExp=1; InAct=EBI-312701, EBI-327326;  
CC O09665:unc-2; NDExp=1; InAct=EBI-312701, EBI-332329;  
CC Q23390:ZK1067.7; NDExp=1; InAct=EBI-312701, EBI-312743;  
CC P34557:ZK632.12; NDExp=1; InAct=EBI-312701, EBI-332273;  
CC EMBL; AL132948; CAC51077.1; -, Genomic\_DNA.  
DR PIR; T45059; T45059.  
DR HSSP; P13231; IHCE.  
DR InAct; OQNE57; -;  
DR Ensemble; Y3986A.1; Caenorhabditis elegans.  
DR Wormbase; WBGene00012664; Y3986A.1.  
DR Wormpep; Y3986A.1; CE21682.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 735 AA; 80254 MW; 6E7B831DBE41159 CRC64;

Query Match 52.6%; Score 50; DB 2; Length 735;  
Best Local Similarity 43.8%; Pred. No. 61;  
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 HLHNVHKEEHAHN 16  
Db 583 HGVDHGHSHSGHGH 598

RESULT 33  
O6YSC5\_ORYSA PRELIMINARY; PRT; 498 AA.  
ID O6YSC5\_ORYSA  
AC O6YSC5;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
DE Putative IAA-alanine resistance protein.  
GN Name=P0104802.41; Synonym=P0470B03.12;  
OS Oryza sativa (Japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzoideae; Oryza.  
OX NCBI\_TaxId=39947;  
RN [1]

RP NUCLEOTIDE SEQUENCE.  
RA Sasaki T., Matsumoto T., Katayose Y.;  
RA Submitted (May-2003) to the EMBL/GenBank/DBJ databases.  
RN [2]

RP NUCLEOTIDE SEQUENCE.  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RA Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP006461; BAD10785.1; -, Genomic\_DNA.  
DR EMBL; AP004585; BAD09550.1; -, Genomic\_DNA.  
DR Gramene; O6YSC5; -;  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0046873; F:metal ion transporter activity; IEA.  
DR GO; GO:0030001; F:metal ion transport; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR003689; Zn\_transp\_Zfp.  
DR Pfam; PF02535; Zfp; 1.  
SQ SEQUENCE 498 AA; 53578 MW; 015101991517C69D CRC64;

Query Match 52.1%; Score 49.5; DB 2; Length 498;  
Best Local Similarity 56.2%; Pred. No. 47;  
Matches 9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

OY 1 HLHNVHKEEHAHN 16  
Db 200 HONHN-HSHSHSHS 214

RESULT 34  
O8J226\_9EURO PRELIMINARY; PRT; 261 AA.  
ID O8J226\_9EURO  
AC O8J226;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Transcription factor AflR (fragment).  
GN Name=AflR;  
OS Aspergillus alliaceus.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Eurotiaceae; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
OX NCBI\_TaxId=209559;  
RN [1]

RP NUCLEOTIDE SEQUENCE.  
RA Cary J.W., Beltz S.B., Montalbano B.G., Ehrlich K.C.;  
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF547172; AAN38761.1; -, Genomic\_DNA.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR002409; Aflatoxin\_BRP.  
DR PRINTS; PR00755; AFLATOXINBRP.  
FT NON\_TER 1 1  
FT NON\_TER 261 261  
SQ SEQUENCE 261 AA; 27174 MW; E1288DF799EBADC9 CRC64;

Query Match 51.6%; Score 49; DB 2; Length 261;  
Best Local Similarity 43.8%; Pred. No. 27;  
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 HLHNVHKEEHAHN 16  
Db 16 HSAHNTYSTRHSHS 31

RESULT 35  
O6NLJ8\_DROME PRELIMINARY; PRT; 374 AA.  
ID O6NLJ8\_DROME  
AC O6NLJ8;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE AT16912P.  
GN Name=Ct14853;

```
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OK NCBI_TaxID=7227;
RN
RP NUCLEOTIDE SEQUENCE.
RA Stapleton M., Carlson J., Chavez C., Friese E., George R., Pacleb J.,
RA Park S., Wan K., Yu C., Rubin G.M., Celniker S.;
RA Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL, F012333; AAS77458.1; -; mRNA.
SQ SEQUENCE 374 AA, 38669 MW, 739C21633B405C70 CRC64;

Query Match 51.6%; Score 49; DB 2; Length 374;
Best Local Similarity 42.9%; Pred. No. 41;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 INNHKEEHAHNN 16
Db 264 VHHSHSHSHSHSH 277

RESULT 36
Q9VFH8_DROME PRELIMINARY; PRT; 374 AA.
AC Q9VFH8; Q8SYG5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE CG14853-PA, Isoform A (CG14853-pb, Isoform b) (BB63320p).
GN ORFName=CG14853; CG14853;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OK NCBI_TaxID=7227;
RN
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballew R.M., Banu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Bugan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pabloe B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostel C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasse K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
RA Jatali M., Kalush F., Karpén G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasbo P., Lei Y., Levitsky A.C., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mirkulov G., Mlshina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiert C., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson R., Skupki M.P., Smith T.,
RA Spitzer E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodege T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RA Science 287:2185-2195(2000).
RN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Friese E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svitek R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RA "Finishing a whole-genome shotgun: release 3 of the Drosophila
RA melanogaster euchromatic genome sequence.";
RA Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svitek R.,
RA Patel S., Friese E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RA "The transposable elements of the Drosophila melanogaster euchromatin:
RA a genomic perspective.";
RA Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Mista S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Belencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RA "Annotation of the Drosophila melanogaster euchromatic genome: a
RA systematic review.";
RA Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN
RP NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Friese E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svitek R., Smith E.,
RA Yu C., Rubin G.;
RA "Drosophila melanogaster release 4 sequence.";
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RA Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Drensek D., Farfan D., Friese E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Pargass V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RA Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN
RP NUCLEOTIDE SEQUENCE.
CC -1- INTERACTION;
CC Q9YV6; CG7386; NDExp=1; InAct=EBI-187640, EBI-151202;
CC Q9YV6; morgue; NDExp=1; InAct=EBI-187640, EBI-89315;
CC CC Q9YH3; UPF3; NDExp=1; InAct=EBI-187640, EBI-16915;
CC EMBL, AE003705; AAF55078.2; -; Genomic_DNA.
CC DR EMBL, AY015172; AAL49194.1; -; mRNA.
CC DR HSSP, Q9Y3D6; 1PC2.
CC DR InAct; Q9VFH8; -;
CC DR Ensembl; CG14853; Drosophila melanogaster.
CC DR FlyBase; Fgn0038246; CG14853.
CC DR SEQUENCE 374 AA, 38729 MW, 91493D5E840AF93A CRC64;

Query Match 51.6%; Score 49; DB 2; Length 374;
```





DT 01-FEB-2005 (TrEMBLrel. 29, last annotation update)  
DE Afir.  
OS Aspergillus nomius.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
OX NCBI\_Taxid=41061;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=AN1337;  
RA Ehrlich K.C.; Chang P.K.; Yu J.; Cotty P.J.;  
RT "Aflatoxin Biosynthesis Cluster Gene cypA Is Required for G Aflatoxin  
RT Formation";  
RL Appl. Environ. Microbiol. 70:6518-6524 (2004).  
DR EMBL; AY510454; AAS90051.1; -; Genomic\_DNA.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0008270; F:zinc ion binding; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR002409; Aflatoxin BRP.  
DR InterPro; IPR001138; Fungi\_Trcrpt\_N.  
DR Pfam; PF00172; Zn\_club; 1.  
DR PRINTS; PR00755; AFLATOXINBRP.  
DR SMART; SM00066; GAL4; 1.  
DR PROSITE; PS00463; ZN2\_CY6\_FUNGAL\_1; 1.  
DR PROSITE; PS50048; ZN2\_CY6\_FUNGAL\_2; 1.  
SQ SEQUENCE 442 AA; 46998 MW; 334CE4B2D6BE7AED CRC64;

Query Match 51.6%; Score 49; DB 2; Length 442;  
Best Local Similarity 37.5%; Pred. No. 49;  
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 HLHHNFKKEHAHNN 16  
DB 93 HSNVNTYSTPWHSHS 108

Search completed: February 11, 2006, 13:14:56  
Job time : 255 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: February 11, 2006, 13:10:58 ; Search time 41 Seconds

(Without alignments)  
37.548 Million cell updates/sec

Title: US-10-762-588-1

Perfect score: 95

Sequence: 1 HLINHYKKEHMAHN 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR 80:\*

1: pirl:\*

2: pirl:\*

3: pirl:\*

4: pirl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	100.0	331	1	DECHLM
-2	95	100.0	332	2	S12151
3	50	52.6	351	1	KGZOHM
4	50	52.6	448	2	E86291
5	50	52.6	735	2	T45059
6	48	50.5	158	2	D36138
7	48	50.5	401	2	T05025
8	48	50.5	498	2	JC1400
9	48	50.5	834	2	G71400
10	48	50.5	1235	2	T16346
11	47	49.5	640	2	A41726
12	47	49.5	642	2	S27606
13	47	49.5	751	2	T15403
14	47	49.5	815	2	T15402
15	46	48.4	325	2	T44365
16	46	48.4	325	2	G90008
17	46	48.4	580	2	T46024
18	46	48.4	580	2	T46024
19	46	48.4	580	2	T46024
20	45.5	47.9	331	1	DECHLM
21	45	47.4	231	2	AD0325
22	45	47.4	231	2	AD0325
23	45	47.4	515	2	T23089
24	45	47.4	946	2	T19465
25	44	46.3	303	2	AG1396
26	44	46.3	303	2	AB1772
27	44	46.3	316	2	E81321
28	44	46.3	338	1	A47183
29	44	46.3	385	2	A84696

30	44	46.3	604	2	A39369	homeotic protein B
31	44	46.3	1226	2	T18459	hypothetical prote
32	43.5	45.8	859	2	S64195	HTRI protein - yea
33	43	45.3	245	2	T04797	embryo-specific pr
34	43	45.3	293	2	T05153	hypothetical prote
35	43	45.3	332	2	A32957	L-lactate dehydrog
36	43	45.3	398	2	T02681	probable zinc tran
37	43	45.3	453	1	T02339	omega-3 fatty acid
38	43	45.3	457	2	S39079	puff C-8 protein -
39	43	45.3	606	2	S13367	Om(1D) protein - f
40	43	45.3	733	2	JC7679	dendritic cell-der
41	43	45.3	744	2	T13429	hypothetical prote
42	42.5	44.7	102	2	T02917	hypothetical prote
43	42.5	44.7	664	2	H70396	cation transportin
44	42	44.2	82	2	A29653	histidine-rich pro
45	42	44.2	157	2	S39849	pfla protein - pse
46	42	44.2	239	2	T07092	Ca+2-binding BP ha
47	42	44.2	387	2	A47446	HNF-3/fork head fa
48	42	44.2	451	2	A55909	transferrin prote
49	42	44.2	459	2	I37451	HBF-G2 (HFK-2) pro
50	42	44.2	476	2	A54743	transcription fact
51	42	44.2	480	2	JH0672	brain factor 1 pro
52	42	44.2	496	2	S33791	ARS-binding protei
53	42	44.2	609	2	F64045	excinuclease ABC c
54	42	44.2	610	2	G82227	excinuclease ABC,
55	42	44.2	1040	2	T29092	TSC-22 protein hom
56	41.5	43.7	143	2	T12245	ABA stress ripenin
57	41.5	43.7	373	2	AD0262	probable membrane
58	41.5	43.7	752	2	G02273	LIV-1 protein - hu
59	41.5	43.7	943	2	T03306	PSD-95/SAP90-asoc
60	41	43.2	242	2	S74235	hyb protein - Syn
61	41	43.2	296	2	G81431	periplasmic solute
62	41	43.2	337	2	AD3614	glycosyl transfera
63	41	43.2	361	2	F87286	cation efflux faml
64	41	43.2	361	2	T19395	hypothetical prote
65	41	43.2	404	2	T25420	hypothetical prote
66	41	43.2	411	2	AG3003	conserved hypotet
67	41	43.2	411	2	B98280	hypothetical prote
68	41	43.2	444	2	F96836	hypothetical prote
69	41	43.2	464	2	G71825	hypothetical prote
70	41	43.2	466	2	C64690	conserved hypotet
71	41	43.2	568	2	S15008	gene disco protein
72	41	43.2	620	2	T30765	hypothetical prote
73	41	43.2	627	2	A99607	hypothetical prote
74	41	43.2	950	2	S27473	URB1 protein - sm
75	41	43.2	1394	2	B34598	ecdysone-induced p
76	41	43.2	2652	1	VF1HB2	genome polyprotein
77	41	43.2	5369	2	T44807	myosinubillin synth
78	40.5	42.6	178	2	T29352	hypothetical prote
79	40.5	42.6	1855	2	S41649	DNA polymerase - m
80	40.5	42.6	2254	2	T09053	low voltage-activa
81	40.5	42.6	2380	2	E71604	hypothetical prote
82	40	42.1	114	2	S37150	asr2 protein - tom
83	40	42.1	131	2	T10490	water-stress-induc
84	40	42.1	143	1	B64421	conserved hypotet
85	40	42.1	149	2	A54530	eggshell protein -
86	40	42.1	175	2	C70326	hypothetical prote
87	40	42.1	211	1	XUBSMM	methylphosphorites
88	40	42.1	237	2	B59107	hypothetical prote
89	40	42.1	251	1	B69978	conserved hypotet
90	40	42.1	294	2	H75331	hypothetical prote
91	40	42.1	306	2	T44684	hypothetical prote
92	40	42.1	315	2	JC7572	somite Maf1 protei
93	40	42.1	320	2	A96570	NAM-like protein,
94	40	42.1	359	2	AF3470	transcription regu
95	40	42.1	385	1	OMRTSP	selonoprotein p pr
96	40	42.1	427	2	A32372	female-specific do
97	40	42.1	453	2	A41640	vestigial protein
98	40	42.1	486	2	T10089	citrin - sweet ora
99	40	42.1	495	1	S31223	transcription fact
100	40	42.1	503	2	S38005	hypothetical prote

## ALIGNMENTS

## RESULT 1

DECHLM

L-lactate dehydrogenase (EC 1.1.1.27) chain M - chicken

C/Species: Gallus gallus (chicken)

C/Date: 18-Aug-1982 #sequence\_revision 18-Aug-1982 #text\_change 09-Jul-2004

C/Accession: A00349

R/Torff, H.J.; Becker, D.; Schwarzwald, J.

In pyridine Nucleotide Dependent Dehydrogenases, Sund, H., ed., pp.31-42, Walter de Gruy

A/Reference number: A94435

A/Accession: A00349

A/Molecule type: protein

A/Residues: 1-331 &lt;PRO&gt;

A/Cross-references: UNIPROT:P00340; UNIPARC:UP10000171F20

C/Function: catalyzes the reversible oxidation of (S)-lactate to pyruvate by NAD+

C/Superfamily: L-lactate dehydrogenase

C/Keywords: acetylated amino end; NAD; oxidoreductase

F1/Modified site: acetylated amino end (Ser) #status predicted

F165.192/Active site: Asp, His #status predicted

Query Match 100.0%; Score 95; DB 1; Length 331;

Best Local Similarity 100.0%; Pred. No. 7.6e-07;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLHNVHKEEHAHNN 16

DB 5 HLHNVHKEEHAHNN 20

## RESULT 2

S12151

L-lactate dehydrogenase (EC 1.1.1.27) chain A - chicken

C/Species: Gallus gallus (chicken)

C/Date: 02-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004

C/Accession: S12151

R/Hirota, Y.; Katsumata, A.; Takeya, T.

Nucleic Acids Res. 18, 6432, 1990

A/Title: Nucleotide and deduced amino acid sequences of chicken lactate dehydrogenase-A.

A/Reference number: S12151; MUID:91057138; PMID:2243792

A/Accession: S12151

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-332 &lt;HR&gt;

A/Cross-references: UNIPROT:P00340; UNIPARC:UP10000171355; EMBL:X53828; NID:963565; PIDN

C/Superfamily: L-lactate dehydrogenase

C/Keywords: oxidoreductase

F166.193/Active site: Asp, His #status predicted

Query Match 100.0%; Score 95; DB 2; Length 332;

Best Local Similarity 100.0%; Pred. No. 7.6e-07;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLHNVHKEEHAHNN 16

DB 6 HLHNVHKEEHAHNN 21

## RESULT 3

KGZQHL

histidine-rich glycoprotein precursor - Plasmodium lophurae

C/Species: Plasmodium lophurae

C/Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 31-Dec-2004

C/Accession: A22692

R/Ravetch, J.V.; Feder, R.; Pavlovic, A.; Blobel, G.

Nature 312, 616-620, 1984

A/Title: Primary structure and genomic organization of the histidine-rich protein of the

A/Reference number: A22692; MUID:85061618; PMID:6095114

A/Accession: A22692

A/Molecule type: DNA

A/Residues: 1-351 &lt;RAV&gt;

A/Cross-references: UNIPROT:P04929; UNIPARC:UP1000012CB80; GB:X01469; NID:99997; PIDN:CA

C/Comment: There are two copies of 16-residue repeats, two copies of 17-residue repeats,

C/Genetics: 23/3

A/Introns: 23/3

C/Superfamily: surface antigen spap

C/Keywords: glycoprotein; tandem repeat

F124-47/Domain: signal sequence #status predicted &lt;SIG&gt;

F124-47/Product: histidine-rich glycoprotein #status predicted &lt;PRO&gt;

F124-47/Region: 16-residue repeats

F124-47/Region: 17-residue repeats

F124-47/Region: 15-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats

F124-47/Region: 10-residue repeats



A;Cross-references: UNIPROT:Q20330; UNIPARC:UPI000017B9B3; EMBL:U40799; NID:g1065935; P  
C;Gene: CBSP:F42C5.10  
A;Introns: 11/3; 280/2; 312/3; 646/1; 786/1; 869/1; 1034/1; 1080/2; 1108/1; 1170/1

Query Match 50.5%; Score 48; DB 2; Length 1225;  
Best Local Similarity 42.9%; Pred. No. 30;  
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 3 HLNHVKEBHAAHN 16  
||:|:|:|:|:|:  
DB 894 VHHYHODENAHQHS 907

RESULT 11  
A41726  
homeotic protein BarH2 - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Dec-2004  
C;Accession: A41726  
R;Highashijima, S.; Kojima, T.; Michue, T.; Ishimaru, S.; Emori, Y.; Saigo, K.  
Gene: Dev. 6, 50-60, 1992  
A;Title: Dual Bar homeo box genes of Drosophila required in two photoreceptor cells, R1  
A;Reference number: A41726; MUID:92112035; PMID:1346120  
A;Accession: A41726  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-640 <HIG>  
A;Cross-references: UNIPROT:Q24256; UNIPARC:UPI00000835EA; GB:M82885; GB:M82886; GB:M828  
A;Note: the authors translated the codon TAC for residue 134 as Thr  
C;Genetics:  
A;Gene: FLYBase:B-H2  
A;Cross-references: FLYBase:Fgn0004854  
C;Keyword: DNA binding; homeobox; nucleus; transcription regulation  
F;376-432/Domain: homeobox homology <HOK>

Query Match 49.5%; Score 47; DB 2; Length 640;  
Best Local Similarity 43.8%; Pred. No. 22;  
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 HLNHVKEBHAAHN 16  
||:|:|:|:|:|:  
DB 235 HLSHLSHQHHPHLHH 250

RESULT 12  
S27806  
homeotic protein BarH2 - fruit fly (Drosophila melanogaster)  
N;Alternate names: dual bar protein  
C;Species: Drosophila melanogaster  
C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 05-Oct-2004  
C;Accession: S27806  
R;Highashijima, S.I.; Kojima, T.; Michue, T.; Ishimaru, S.; Emori, Y.; Saigo, K.  
submitted to the EMBL Data Library, March 1992  
A;Description: Dual bar homeobox genes of Drosophila required two photoreceptor cells, R  
A;Reference number: S27806  
A;Accession: S27806  
A;Molecule type: mRNA  
A;Residues: 1-642 <HIG>  
A;Cross-references: UNIPARC:UPI000017A2E6; EMBL:M82884; NID:g156982; PID:g156983  
C;Genetics:  
A;Gene: FLYBase:B-H2  
A;Cross-references: FLYBase:Fgn0004854  
C;Keyword: DNA binding; homeobox; nucleus; transcription regulation  
F;378-434/Domain: homeobox homology <HOK>

Query Match 49.5%; Score 47; DB 2; Length 642;  
Best Local Similarity 43.8%; Pred. No. 22;  
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 HLNHVKEBHAAHN 16  
||:|:|:|:|:|:  
DB 235 HLSHLSHQHHPHLHH 250

RESULT 13  
T15403  
hypothetical protein C04A2.7 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 23-Mar-2001  
C;Accession: T15403  
R;Du, Z.  
submitted to the EMBL Data Library, July 1995  
A;Description: The sequence of C. elegans cosmid C04A2.  
A;Reference number: S59416  
A;Accession: T15403  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-751 <DUZ>  
A;Cross-references: UNIPARC:UPI000017B728; EMBL:U23448; NID:g733539; PID:g733543; PIDN:A  
A;Experimental source: strain Bristol N2  
C;Genetics:  
A;Gene: CBSP:C04A2.7  
A;Introns: 6/2; 73/3; 129/3; 181/2; 259/3; 382/1; 443/1; 593/3; 632/3; 688/3; 722/3; 746

Query Match 49.5%; Score 47; DB 2; Length 751;  
Best Local Similarity 50.0%; Pred. No. 26;  
Matches 8; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 1 HLNHVKEBHAAHN 16  
||:|:|:|:|:|:  
DB 115 HPHNMLR--HHTHTHN 128

RESULT 14  
T15402  
hypothetical protein C04A2.7a - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 23-Mar-2001  
C;Accession: T15402  
R;Du, Z.  
submitted to the EMBL Data Library, July 1995  
A;Description: The sequence of C. elegans cosmid C04A2.  
A;Reference number: S59416  
A;Accession: T15402  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-815 <DUZ>  
A;Cross-references: UNIPARC:UPI000017B729; EMBL:U23448; NID:g733539; PID:g733544; PIDN:A  
A;Experimental source: strain Bristol N2  
C;Genetics:  
A;Gene: CBSP:C04A2.7a  
A;Introns: 6/2; 73/3; 129/3; 181/2; 259/3; 382/1; 443/1; 593/3; 632/3; 688/3; 722/3; 746

Query Match 49.5%; Score 47; DB 2; Length 815;  
Best Local Similarity 50.0%; Pred. No. 28;  
Matches 8; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 1 HLNHVKEBHAAHN 16  
||:|:|:|:|:|:  
DB 115 HPHNMLR--HHTHTHN 128

RESULT 15  
T44365  
cation-efflux system membrane protein homolog (imported) - Staphylococcus aureus  
C;Species: Staphylococcus aureus  
C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
C;Accession: T44365  
R;Kuroda, M.; Hayashi, H.; Ohta, T.  
Microbiol. Immunol. 43, 115-125, 1999  
A;Title: Chromosome-determined zinc responsive operon czr in Staphylococcus aureus stra  
A;Reference number: Z2734; MUID:99244271; PMID:10228265  
A;Accession: T44365  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA

FEBS Lett. 321, 93-97, 1993



A>Title: Cloning, expression and characterization of human kininogen domain 3.  
A:Reference number: S32422; MUID:93223854; PMID:8467916  
A:Accession: S32422  
A:Molecule type: mRNA  
A:Residues: 'ANSM', 253-377 <AUE>  
A:Cross-references: UNIPARC:UPI0000174297  
A>Note: differences are due to known cloning artifacts  
R:Loetspeich, F.; Kellermann, J.; Henschen, A.; Peetersch, B.; Muller-Esterl, W.  
Eur. J. Biochem. 152, 307-314, 1985  
A>Title: The amino acid sequence of the light chain of human high-molecular-mass kininogen  
A:Reference number: A91153; MUID:86030270; PMID:4054110  
A:Accession: A91153  
A:Molecule type: protein  
A:Residues: 379-644 <LOT>  
A:Cross-references: UNIPARC:UPI0000174298  
A>Note: the bradykinin sequence preceding the light chain sequence was not determined in R:Kellermann, J.; Loetspeich, F.; Henschen, A.; Mueller-Esterl, W.  
Eur. J. Biochem. 154, 471-478, 1986  
A>Title: Completion of the primary structure of human high-molecular-mass kininogen. The A:Reference number: A24871; MUID:86108361; PMID:3484703  
A:Accession: A24871  
A:Molecule type: protein  
A:Residues: 'Z', 20-380 <KEL1>  
A:Cross-references: UNIPARC:UPI0000174299  
R:Kellermann, J.; Loetspeich, F.; Henschen, A.; Mueller-Esterl, W.  
in: Kinins IV, Greenbaum, L.M., and Margolis, H.S., ed., pp.85-89, Plenum Press, New York  
A>Title: Amino acid sequence of the light chain of human high molecular mass kininogen.  
A:Reference number: A27899  
A:Accession: A27899  
A:Molecule type: protein  
A:Residues: 379-389, 'K', 390-407, 'Q', 409-644 <KEL2>  
A:Cross-references: UNIPARC:UPI000017429A  
R:Mindroiu, T.; Carretero, O.A.; Proud, D.; Malz, D.; Scicl, A.G.  
Biochem. Biophys. Res. Commun. 152, 519-526, 1988  
A>Title: A new kinin moiety in human plasma kininogens.  
A:Reference number: A27699; MUID:88209021; PMID:3365237  
A:Accession: A27699  
A:Molecule type: protein  
A:Residues: 380-389 <MIN>  
A:Cross-references: UNIPARC:UPI000002C078  
R:Maeda, H.; Matsumura, Y.; Kato, H.  
J. Biol. Chem. 263, 16051-16054, 1988  
A>Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic fluid  
A:Reference number: A31905; MUID:89034061; PMID:3182782  
A:Accession: A31905  
A:Molecule type: protein  
A:Residues: 381-389 <MAB>  
A:Cross-references: UNIPARC:UPI000002CF4A  
R:Saeguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.  
Biochem. Biophys. Res. Commun. 150, 511-516, 1988  
A>Title: Identification of [hydroxyproline(3)]-tyrosyl-bradykinin released from human plasma  
A:Reference number: A34030; MUID:88106652; PMID:3337729  
A:Accession: A34030  
A:Molecule type: protein  
A:Residues: 380-389 <ASAS>  
A:Cross-references: UNIPARC:UPI000002C078  
R:Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.  
Biol. Chem. Hoppe-Seyler 369, 257-261, 1988  
A>Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory and A:Reference number: S02482; MUID:89076517; PMID:3264507  
A:Accession: S02482  
A:Molecule type: protein  
A:Residues: 1-19,189-192;310-314;381-389 <LEN1>  
A:Cross-references: UNIPARC:UPI000002CF4A; UNIPARC:UPI0000143284; UNIPARC:UPI000017429B; R:Kato, H.; Matsumura, Y.; Maeda, H.  
FEBS Lett. 232, 252-254, 1988  
A>Title: Isolation and identification of hydroxyproline analogues of bradykinin in human A:Reference number: A61495; MUID:88211869; PMID:3366244  
A:Accession: A61495  
A:Molecule type: protein  
A:Residues: 380-389 <KAT1>  
A:Cross-references: UNIPARC:UPI000002C078  
A:Experimental source: urine

A>Note: this peptide had Pro-383 modified to 4-hydroxyproline  
A:Accession: B61495  
A:Molecule type: protein  
A:Residues: 381-389 <KAT2>  
A:Cross-references: UNIPARC:UPI000002CF4A  
A:Experimental source: urine  
A>Note: this peptide had Pro-383 modified to 4-hydroxyproline  
A:Accession: C61495  
A:Molecule type: protein  
A:Residues: 380-389 <KAT3>  
A:Cross-references: UNIPARC:UPI000002C078  
R:Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.  
FEBS Lett. 280, 211-215, 1991  
A>Title: Inactivation of human cystatin C and kininogen by human cathepsin D.  
A:Reference number: S14303; MUID:91192133; PMID:2013314  
A:Accession: S14447  
A:Molecule type: protein  
A:Residues: 264-359, 'N', 361-375 <LEN2>  
A:Cross-references: UNIPARC:UPI000017429D  
R:Little, S.S.; Johnson, D.A.  
Biochem. J. 307, 341-346, 1995  
A>Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec A:Reference number: S55239; MUID:95251593; PMID:7733867  
A:Accession: S55239  
A:Molecule type: protein  
A:Residues: 450-452, 'X', 454, 'X', 456 <LIT>  
A:Cross-references: UNIPARC:UPI000017429E  
R:Straczek, J.; Maachi, F.; Je Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabec, P.; Bellevil  
FEBS Lett. 373, 207-211, 1995  
A>Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like A:Reference number: S68059; MUID:96033974; PMID:7589467  
A:Accession: S68059  
A:Molecule type: protein  
A:Residues: 431-434 <STR>  
A:Cross-references: UNIPARC:UPI000017429F  
R:Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.  
J. Biol. Chem. 260, 8610-8617, 1985  
A>Title: Structural organization of the human kininogen gene and a model for its evolution  
A:Reference number: A92545; MUID:85234583; PMID:2989294  
A:Accession: A92545  
A:Contents: annotation; gene organization  
R:Pierce, J.V.  
Fed. Proc. 27, 52-57, 1968  
A>Title: Structural features of plasma kinins and kininogens.  
A:Reference number: A91455; MUID:90255622; PMID:4952632  
A:Contents: annotation; bradykinin  
C:Comment: The HMW kininogen precursor and the LMW form are produced from the same gene  
C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is important  
C:Comment: Kininogen, released from kininogen by kallikrein, is a potent vasodilator, 1  
xyproline residue is present in the kininogen prior to the release of bradykinin.  
C:Genetics:  
A:Gene: GDB:KNG  
A:Cross-references: GDB:125256; OMIM:228960  
A:Map position: 3q27-3q27  
A:Intons: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3  
C:Superfamily: Kininogen; cystatin homology  
C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl  
F:1-16/Domain: signal sequence #status experimental <STG>  
F:19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>  
F:19-379/Domain: HMW kininogen II #status experimental <MAT2>  
F:19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>  
F:19-331/Domain: cystatin homology <CV1>  
F:142-253/Domain: cystatin homology <CV2>  
F:1264-375/Domain: cystatin homology <CV3>  
F:380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDV>  
F:381-389/Product: bradykinin (kallidin I) #status experimental <BDV>  
F:390-644/Domain: HMW kininogen light chain #status experimental <LCH>  
F:421-510/Region: glycine/histidine/lysine-rich 30-residue repeats  
F:431-434/Product: low molecular weight growth promoting factor #status experimental <GF  
F:19/Modified site: pyroglutamate carboxylic acid (Gln) (in mature form) #status experimen  
F:28-614,83-94,107-126,142-145,206-214,229-246,264-267,328-340,351-370/Diulfide bonds:  
F:48/Binding site: carboxylate (Asn) (covalent) #status absent  
F:169,205,294/Binding site: carboxylate (Asn) (covalent) #status experimental



F;379-380/Cleavage site: Met-Lys (kallikrein) #status experimental  
F;383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental  
F;389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental  
F;401,533,542,546,557,571,628/Binding site: carbohydrate (Thr) (covalent) #status experimental  
F;577/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 48.4%; Score 46; DB 1; Length 644;  
Best Local Similarity 40.0%; Pred. No. 31;  
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 HLHNHKEHNAH 15  
DB 489 HVLDDHGKHHGH 503

RESULT 20  
DEPGIM  
L-lactate dehydrogenase (EC 1.1.1.27) chain M - pig  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 07-May-1981 #sequence\_revision 07-May-1981 #text\_change 09-Jul-2004  
C;Accession: A00348  
R;Kilby, H.H.; Keil, W.; Griesbach, M.; Petry, K.; Meyer, H.  
Hoppe-Seyler's Z. Physiol. Chem. 358, 123-127, 1977  
A;Title: The primary structure of porcine lactate dehydrogenase: isoenzymes M-4 and H-4.  
A;Reference number: A91671; PMID:77117453; PMID:838465  
A;Accession: A00348  
A;Molecule type: protein  
A;Residues: 1-331 <KIL>  
A;Cross-references: UNIPROT:P00339; UNIPARC:UPI0000171F1F  
C;Function:  
A;Description: catalyzes the reversible oxidation of (S)-lactate to pyruvate by NAD+  
C;Superfamily: L-lactate dehydrogenase  
C;Keywords: acetylated amino end; NAD; oxidoreductase; tetramer  
F;1/Modified site: acetylated amino end (Ala) #status experimental  
F;165,192/Active site: Asp, His #status predicted

Query Match 47.9%; Score 45.5; DB 1; Length 331;  
Best Local Similarity 66.7%; Pred. No. 18;  
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 2 LHNHKEHNAH 16  
DB 7 LHNHKEHNAH-PHN 20

RESULT 21  
AD0325  
urease accessory protein [imported] - Yersinia pestis (strain CO92)  
C;Species: Yersinia pestis  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 31-Dec-2004  
C;Accession: AD0325  
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.  
R;Farrell, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.;  
Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,  
Nature 413, 523-527, 2001  
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A;Reference number: AB0001; PMID:11586360  
A;Accession: AD0325  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-231 <KUR>  
A;Cross-references: UNIPROT:Q9ZFR8; UNIPARC:UPI0000137DC6; GB:AL590842; PIDN:CAC92907.1;  
C;Genetic:  
A;Gene: ureB  
C;Superfamily: urease accessory protein UreB (nickel metallochaperone)

Query Match 47.4%; Score 45; DB 2; Length 231;  
Best Local Similarity 46.2%; Pred. No. 15;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 HNHNKEHNAH 15  
DB 199 HNHNKEHNAH 211

RESULT 22  
D83483  
probable metal transporter PA1297 [imported] - Pseudomonas aeruginosa (strain PA01)  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C;Accession: D83483  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Muzoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B  
adman, S.; Ryan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim  
, J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path  
A;Reference number: A82950; PMID:10437337; PMID:10984043  
A;Accession: D83483  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-326 <STO>  
A;Cross-references: UNIPROT:Q91447; UNIPARC:UPI00000C52F8; GB:AE004559; GB:AE004091; NI  
C;Genetic:  
A;Gene: PA1297  
C;Superfamily: zinc transporter Znt-2

Query Match 47.4%; Score 45; DB 2; Length 326;  
Best Local Similarity 43.8%; Pred. No. 21;  
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 HLHNHKEHNAH 16  
DB 168 HNHNHNDHNAHRHD 183

RESULT 23  
T23089  
hypothetical protein H13N06.5 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T23089  
R;Jennard, N.  
submitted to the EMBL Data Library, October 1997  
A;Reference number: Z19673  
A;Accession: T23089  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-515 <WIL>  
A;Cross-references: UNIPROT:Q9XT07; UNIPARC:UPI000046CA2; EMBL:Z99942; PIDN:CAB17070.1,  
A;Experimental source: clone H13N06  
C;Genetic:  
A;Gene: CBSP:H13N06.5  
A;Map position: X  
A;Introns: 118/1; 156/2; 182/2; 306/1

Query Match 47.4%; Score 45; DB 2; Length 515;  
Best Local Similarity 63.6%; Pred. No. 34;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 NVNKEHNAH 15  
DB 90 NVNKEHNAH 100

RESULT 24  
T19465  
hypothetical protein M04C3.1 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T19465; T23703  
R;Northmore, B.  
submitted to the EMBL Data Library, November 1996  
A;Reference number: Z19126  
A;Accession: T19465  
A;Status: preliminary; translated from GB/EMBL/DBJ



A/Title: Expression, characterization, and crystallization of oxygen-avid Ascaris hemoglobin  
A/Reference number: A47403; MUID:93352567; PMID:8349648  
A/Contents: annotation; physical properties  
R/de Beere, I.; Liu, L.; Moens, L.; Van Beuningen, J.; Gielens, C.; Richelle, J.; Trotman, P.  
Proc. Natl. Acad. Sci. U.S.A. 89, 4638-4642, 1992  
A/Title: Polar zipper sequence in the high-affinity hemoglobin of Ascaris suum: amino acid  
A/Reference number: A45289; MUID:9252495; PMID:1564800  
A/Accession: A45289  
A/Molecule type: protein  
A/Residues: 19-116, 'L', 118-168, 173-265, 'D', 267-333 <DEA>  
A/Cross-references: UNIPARC:UPI000017391A  
C/Comment: This extracellular hemoglobin is an octamer of identical chains. It has an ex  
C/Complex: homooctamer  
C/Superfamily: globin, nematode type; globin homology  
C/Keywords: chromoprotein; duplication; extracellular protein; glycoprotein; heme; homoc  
F/1-18/Domain: signal sequence #status predicted <SIG>  
F/22-167/Domain: globin homology <GLB2>  
F/171-316/Domain: globin homology <GLB2>  
F/317-336/Region: 4-residue repeats (E-H-K-E)  
F/19,216/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F/82/Binding site: oxygen (Gln) (distal axial ligand) #status predicted  
F/114/Binding site: heme iron (His) (proximal axial ligand) #status predicted  
F/231/Binding site: oxygen (Gln) (distal axial ligand) #status predicted  
F/263/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 46.3%; Score 44; DB 1; Length 338;  
Best Local Similarity 58.3%; Pred. No. 31;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

4 HNVHKEHNAH 15  
318 HNEHKEHKEH 329

Db HNEHKEHKEH 329

RESULT 29  
A/Title: probable zinc transporter [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: A84696  
R/Jin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;  
eues, D.; Niemman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, L.  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MUID:20083487; PMID:10617197  
A/Accession: A84696  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-385 <STO>  
A/Cross-references: UNIPROT:Q92W23; UNIPARC:UPI00000A37C5; GB:AE002093; NID:G3980394; PI  
C/Genetics:  
A/Gene: AT2G29410  
A/Map position: 2

Query Match 46.3%; Score 44; DB 2; Length 385;  
Best Local Similarity 46.7%; Pred. No. 36;  
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

1 HLHNHKEHNAH 15  
202 HHHHHNHKHQHNNH 216

Db HHHHHNHKHQHNNH 216

RESULT 30  
A/Title: homeotic protein BarH1 - fruit fly (Drosophila ananassae)  
C/Species: Drosophila ananassae  
C/Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 31-Dec-2004  
C/Accession: A39369  
R/Kojima, T.; Ishimaru, S.; Higashijima, S.; Takayama, E.; Akimaru, H.; Sone, M.; Emori,  
Proc. Natl. Acad. Sci. U.S.A. 88, 4333-4347, 1991  
A/Title: Identification of a different-type homeobox gene, BarH1, possibly causing Bar

A/Reference number: A39369; MUID:91239564; PMID:1674606  
A/Accession: A39369  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-604 <KOJ>  
A/Cross-references: UNIPROT:Q23838; UNIPARC:UPI0000075A37; GB:M59962; GB:M59963; NID:gl  
C/Genetics:  
A/Gene: FlyBase:Dana/B  
A/Cross-references: FlyBase:FBgn0012114  
C/Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F/331-387/Domain: homeobox homology <HOX>

Query Match 46.3%; Score 44; DB 2; Length 604;  
Best Local Similarity 35.7%; Pred. No. 57;  
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

3 HNVHKEHNAH 16  
165 LNHLPQSHPHPS 178

Db LNHLPQSHPHPS 178

RESULT 31  
A/Title: hypothetical protein C0515C - malaria parasite (Plasmodium falciparum)  
C/Species: Plasmodium falciparum  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T18459  
R/Lawson, D.; Bowman, S.; Barrell, B.  
submitted to the EMBL Data Library, November 1998  
A/Reference number: Z18937  
A/Accession: T18459  
A/Status: preliminary; translated from GB/EMBL/DDBJ  
A/Molecule type: DNA  
A/Residues: 1-1236 <LAW>  
A/Cross-references: UNIPROT:O77346; UNIPARC:UPI000007F8A1; EMBL:AL008970; NID:e1407852;  
C/Genetics:  
A/Map position: 3  
A/Note: C0515C

Query Match 46.3%; Score 44; DB 2; Length 1236;  
Best Local Similarity 87.5%; Pred. No. 1,2e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

3 HNVHKEH 10  
1166 IHNHKEH 1173

Db IHNHKEH 1173

RESULT 32  
A/Title: HTPI protein - yeast (Saccharomyces cerevisiae)  
N/Alternate names: protein B1834; protein G1625; protein YGL178w; suppressor protein M  
C/Species: Saccharomyces cerevisiae  
C/Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 09-Jul-2004  
C/Accession: S64195; S51561; S50194; S57253; S56159; S50156  
R/Brusch, C.V.; Coglieva, M.; Betanai, I.; Klima, R.; Zaccaria, P.; Delneri, D.  
submitted to the Protein Sequence Database, May 1996  
A/Reference number: S64183  
A/Accession: S64195  
A/Molecule type: DNA  
A/Residues: 1-859 <BRU>  
A/Cross-references: UNIPROT:P39016; UNIPARC:UPI0000168329; EMBL:Z72700; NID:g1322785; P  
A/Experimental source: strain S288C  
R/Kikuchi, Y.; Oka, Y.; Kobayashi, M.; Uesono, Y.; Toh-e, A.; Kikuchi, A.  
Mol. Gen. Genet. 245, 107-116, 1994  
A/Title: A new yeast gene, HTPI, required for growth at high temperature, is needed for  
A/Reference number: S51561; MUID:95147841; PMID:7845552  
A/Accession: S51561  
A/Molecule type: DNA  
A/Residues: 26-859 <KIR>  
A/Cross-references: UNIPARC:UPI000012F4B8; EMBL:D25541; NID:G436256; PIDN:BA05024.1; P  
R/Sakai, A.  
submitted to the EMBL Data Library, December 1993

A/Description: Multicopy suppressors of the yeast pop2 mutation.  
A/Reference number: S48511  
A/Accession: S50194  
A/Molecule type: DNA  
A/Residues: 26-859 <SAS>  
A/Cross-references: UNIPARC:UPI000012F4B8; EMBL:D26184; NID:G450489; PIDN:BAA05172.1; PI  
R:Coglievina, M.; Bartani, I.; Klima, R.; Zaccaria, P.; Brusch, C.V.  
Yeast 11, 767-774, 1995  
A/Title: The DNA sequence of a 7941 bp fragment of the left arm of chromosome VII of Sac  
charomyces cerevisiae and a putative serine/threonine protein kinase gene.  
A/Reference number: S57252; MIMD:95357594; PMID:768046  
A/Accession: S57253  
A/Status: nucleic acid sequence not shown  
A/Molecule type: DNA  
A/Residues: 26-859 <COG>  
A/Cross-references: UNIPARC:UPI000012F4B8; EMBL:X83690; NID:G794143; PIDN:CAA58663.1; PI  
R:Brusch, C.V.  
Submitted to the EMBL Data Library, January 1995  
A/Reference number: S56168  
A/Accession: S56169  
A/Molecule type: DNA  
A/Residues: 1-801, 'Q', 812, 'NCTPRRIML' <BRW>  
A/Cross-references: UNIPARC:UPI0000168BD; EMBL:X83690; NID:G794143; PIDN:CAA58660.1; PI  
C/Genetics:  
A/Gene: SGD:MPT5; HTR1  
A/Cross-references: SGD:S0003146; MIPS:YGL178w  
A/Map position: 7L  
A/Introns: 1/3  
A/Keywords: transmembrane protein  
C/Keywords: transmembrane #status predicted <TM>  
F:497-513/Domain: transmembrane #status predicted <TM>  
  
Query Match 45.8%; Score 43.5; DB 2; Length 859;  
Best Local Similarity 61.5%; Pred. No. 98;  
Matches 8; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
  
QY 4 HNNHKEHHAAH 16  
DB 636 HNNHKEHHAAH 647  
  
RESULT 33  
T04797  
embryo-specific protein 1 - Arabidopsis thaliana  
N/Alternate names: protein F10M23.80  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C/Accession: T04797  
R:Bevan, M.; Lecharny, A.; Chedford, F.; Krivitzky, M.; Kreisel, M.; Hohnsbeil, J.; Mewes, H  
submitted to the Protein Sequence Database, February 1999  
A/Reference number: Z15385  
A/Accession: T04797  
A/Molecule type: DNA  
A/Residues: 1-245 <BEV>  
A/Cross-references: UNIPROT:O81270; UNIPARC:UPI000009E054; EMBL:AL035440  
A/Experimental source: cultivar Columbia; BAC clone F10M23  
C/Genetics:  
A/Map position: 4  
A/Introns: 40/1, 90/1, 118/3, 150/2, 192/2  
A/Note: F10M23.80  
C/Superfamily: rice abscisic acid-induced protein  
  
Query Match 45.3%; Score 43; DB 2; Length 245;  
Best Local Similarity 66.7%; Pred. No. 31;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 HNNHKEHH 11  
DB 130 HNNHKEHH 138  
  
RESULT 34  
T05153

hypothetical protein F18E5.70 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C/Accession: T05153  
R:Bevan, M.; Peter, S.A.; van Staveren, M.; Dirks, W.; Stiekema, W.; Bancroft, I.; Ne  
submitted to the Protein Sequence Database, August 1998  
A/Reference number: Z15400  
A/Accession: T05153  
A/Molecule type: DNA  
A/Residues: 1-293 <BEV>  
A/Cross-references: UNIPROT:O65410; UNIPARC:UPI0000049C90; EMBL:AL022603  
A/Experimental source: cultivar Columbia; BAC clone F18E5  
C/Genetics:  
A/Map position: 4  
A/Introns: 120/1, 143/3, 180/3, 208/3, 234/3, 274/3  
A/Note: F18E5.70  
  
Query Match 45.3%; Score 43; DB 2; Length 293;  
Best Local Similarity 46.2%; Pred. No. 38;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
  
QY 3 HNNHKEHHAAH 15  
DB 54 HNNHKEHHAAH 66  
  
RESULT 35  
A32957  
L-lactate dehydrogenase (EC 1.1.1.27) chain M - rabbit  
C/Species: Oryctolagus cuniculus (domestic rabbit)  
C/Date: 22-Nov-1989 #sequence\_revision 04-Sep-1992 #text\_change 09-Jul-2004  
C/Accession: A32957  
R:Saeg, C.; Briand, M.; Benslimane, S.; Renaud, M.; Briand, Y.  
U. Biol. Chem. 264, 4076-4081, 1989  
A/Title: Characterization of rabbit lactate dehydrogenase-M and lactate dehydrogenase-H  
A/Reference number: A32957; MIMD:89139477; PMID:2917988  
A/Accession: A32957  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-332 <SAS>  
A/Cross-references: UNIPROT:P13491; UNIPARC:UPI000016C581; GB:M2585; NID:G165452; PIDN  
A/Note: the authors translated the codon TGC for residue 163 as Lys  
C/Superfamily: L-lactate dehydrogenase  
C/Keywords: NAD; oxido-reductase  
F:166,193/Active site: Asp, His #status predicted  
  
Query Match 45.3%; Score 43; DB 2; Length 332;  
Best Local Similarity 80.0%; Pred. No. 43;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 LHHNHKEHH 11  
DB 8 LHHNHKEHH 17  
  
RESULT 36  
T02681  
probable zinc transporter At2g46800 (imported) - Arabidopsis thaliana  
N/Alternate names: hypothetical protein F19D11.8  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C/Accession: T02681; D84907  
R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Ron  
submitted to the EMBL Data Library, September 1998  
A/Description: Arabidopsis thaliana chromosome II BAC F19D11 genomic sequence.  
A/Reference number: Z14698  
A/Accession: T02681  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-398 <ROU>  
A/Cross-references: UNIPROT:O81036; UNIPARC:UPI000009C8B0; EMBL:AC005310; NID:G3510247;  
A/Experimental source: cultivar Columbia  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402:761-768, 1999

A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: D84907

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1398 <STO>

A:Cross-references: UNIPARC:UPI000009C8B0; GB:AE020293; NID:93510254; PIDN:AMC3498.1; C:Genetics:

A:Gene: At2g46800; F19D11.8

A:Map position: 2

Query Match 45.3%; Score 43; DB 2; Length 398;  
Best Local Similarity 50.0%; Pred. No. 52;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 4 HNVKKEHAAHN 15  
DB 211 HHHHDEHGHSH 222

RESULT 37

omega-3 fatty acid desaturase (EC 1.14.99.-) GMD [similarity] - soybean

C:Species: Glycine max (soybean)

C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 05-Oct-2004

C:Accession: J02339

R:Yadav, N.S.; Wierzbicki, A.; Aegerter, M.; Caster, C.S.; Perez-Grau, L.; Kinney, A.J.; J. Russell, S.H.; Feldmann, K.A.; Pierce, J.; Browse, J.

Plant Physiol. 103, 467-476, 1993

A>Title: Cloning of higher plant omega-3 fatty acid desaturases.

A:Reference number: J02335; MUID:94302147; PMID:8029334

A:Contents: cDNA:GMD

A:Accession: J02339

A:Molecule type: mRNA

A:Residues: 1453 <YAD>

A:Cross-references: UNIPROT:P48621; UNIPARC:UPI000012A5C; GB:I22965; NID:9408791; PIDN: C:Comment: This enzyme introduces the third double bond in the biosynthesis of 18:2 and C:Superfamily: fatty acid (acyl-CoA) desaturase

C:Keywords: chloroplast; oxidoreductase

Query Match 45.3%; Score 43; DB 1; Length 453;  
Best Local Similarity 46.2%; Pred. No. 59;  
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 4 HNVKKEHAAHN 16  
DB 207 HRTTHQHGHAAHN 219

RESULT 38

S39079

puft C-8 protein - fungus gnat (*Rhynchosciara americana*)

C:Species: *Rhynchosciara americana*

C>Date: 25-Oct-1994 #sequence\_revision 01-Dec-1995 #text\_change 09-Jul-2004

C:Accession: S39079

R:Frydman, H.M.; Cadavid, E.O.; Yokosawa, J.; Silva, F.H.; Navarro-Cattapan, L.D.; Sante J. Mol. Biol. 233, 799-803, 1993

A>Title: Molecular characterization of the DNA puff C-8 gene of *Rhynchosciara americana*.

A:Reference number: S39079; MUID:94016595; PMID:8411183

A:Accession: S39079

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1457 <FRY>

A:Cross-references: UNIPROT:Q26227; UNIPARC:UPI0000084009; EMBL:X64590; NID:9414685; PID

Query Match 45.3%; Score 43; DB 2; Length 457;  
Best Local Similarity 53.8%; Pred. No. 60;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 4 HNVKKEHAAHN 16

DB 55 HSRHNRCHAAHNN 67

RESULT 39

S13367

Om(1D) protein - fruit fly (*Drosophila ananassae*)

C:Species: *Drosophila ananassae*

C>Date: 18-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 05-Oct-2004

C:Accession: S13367

R:Randa, S.; Corcos, V.G. EMBO J. 10, 407-417, 1991

A>Title: Retrotransposon-induced overexpression of a homeobox gene causes defects in ey

A:Reference number: S13367; MUID:9112048; PMID:1671353

A:Accession: S13367

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1606 <RAN>

A:Cross-references: UNIPROT:P22544; UNIPARC:UPI000012C97A; EMBL:X56682; NID:97146; PIDN: C:Genetics:

A:Gene: FlyBase:Dana/B

A:Cross-references: FlyBase:FBgn0012114

A:Initons: 307/1; 383/2

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:332-388/Domain: homeobox homology <Hox>

Query Match 45.3%; Score 43; DB 2; Length 606;  
Best Local Similarity 38.5%; Pred. No. 81;  
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 3 HNVKKEHAAHN 15  
DB 165 LHHHMQSHHPRH 177

RESULT 40

JC7679

dendritic cell-derived BTB/POZ zinc finger protein, DPZF - human

C:Species: Homo sapiens (man)

C>Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001

C:Accession: JC7679

R:Zhang, W.; Li, J.; Li, N.; Su, L.; Wan, T.; Zhang, J.; Chen, T.; Cao, X. Biochem. Biophys. Res. Commun. 282, 1067-1073, 2001

A>Title: Identification and characterization of DPZF, a novel human BTB/POZ zinc finger

A:Reference number: JC7679; MUID:21250664; PMID:11352661

A:Accession: JC7679

A:Molecule type: mRNA

A:Residues: 1733 <ZHA>

A:Cross-references: UNIPARC:UPI000017CB63; GB:AF139460

A:Experimental source: Dendritic cells

C:Comment: This protein, belonging to the subfamily of BTB/POZ zinc finger proteins, isogenesis, lymphoid development, and in immune responses.

C:Genetics:

A:Gene: dpzf

A:Map position: 3q27

C:Keywords: transcription factor; zinc finger

F:573-686/Domain: four tandemly repeated C2H2 zinc fingers (carboxyl terminus) #status :

Query Match 45.3%; Score 43; DB 2; Length 733;  
Best Local Similarity 50.0%; Pred. No. 98;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 LHHVKEHAAHN 15  
DB 48 LHHSTSLTNSHAAH 61

Search completed: February 11, 2006, 13:15:43  
Job time : 44 secs

This Page Blank (Intentional)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2006, 13:16:48 ; Search time 17 Seconds  
(without alignments)  
12.351 Million cell updates/sec

Title: US-10-762-588-1

Perfect score: 95  
Sequence: 1 HLHNVHKEHMAHN 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 97014 seqs, 1312538 residues

Total number of hits satisfying chosen parameters: 97014

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA.New:\*  
1: /cgn2\_6/prodata/1/pubppa/US06\_NEW\_PUB.pcp:\*  
2: /cgn2\_6/prodata/1/pubppa/US07\_NEW\_PUB.pcp:\*  
3: /cgn2\_6/prodata/1/pubppa/PCF\_NEW\_PUB.pcp:\*  
4: /cgn2\_6/prodata/1/pubppa/US09\_NEW\_PUB.pcp:\*  
5: /cgn2\_6/prodata/1/pubppa/US10\_NEW\_PUB.pcp:\*  
6: /cgn2\_6/prodata/1/pubppa/US11\_NEW\_PUB.pcp:\*  
7: /cgn2\_6/prodata/1/pubppa/US60\_NEW\_PUB.pcp:\*  
8: /cgn2\_6/prodata/1/pubppa/US60\_NEW\_PUB.pcp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47.5	50.0	248	US-11-098-686-10847	Sequence 10847, A
2	43	45.3	29	US-11-108-185-51	Sequence 51, Appl
3	42	44.2	274	US-11-015-546A-12	Sequence 12, Appl
4	42	44.2	481	US-11-116-939-14	Sequence 14, Appl
5	42	44.2	696	US-11-242-243-2	Sequence 2, Appl
6	42	44.2	829	US-10-512-109-27	Sequence 27, Appl
7	42	44.2	832	US-10-512-109-29	Sequence 29, Appl
8	41.5	43.7	32	US-11-054-281-104	Sequence 104, Appl
9	41.5	43.7	749	US-11-054-281-103	Sequence 103, Appl
10	41.5	43.7	752	US-11-080-991-52	Sequence 102, Appl
11	41.5	43.7	752	US-11-054-281-101	Sequence 101, Appl
12	41.5	43.7	755	US-10-453-372-1108	Sequence 1108, Ap
13	41.5	43.7	755	US-11-054-281-26	Sequence 26, Appl
14	41.5	43.7	755	US-11-054-281-102	Sequence 102, Appl
15	41	43.2	16	US-10-968-790-27	Sequence 27, Appl
16	40	42.1	469	US-11-124-368A-336	Sequence 336, Appl
17	40	42.1	469	US-11-124-368A-337	Sequence 337, Appl
18	39	41.1	21	US-11-185-111-41	Sequence 41, Appl
19	39	41.1	21	US-11-148-108-51	Sequence 51, Appl
20	39	41.1	21	US-11-094-071-31	Sequence 31, Appl
21	39	41.1	29	US-11-108-185-49	Sequence 49, Appl
22	39	41.1	29	US-11-108-185-52	Sequence 52, Appl
23	39	41.1	98	US-10-952-535A-15	Sequence 15, Appl
24	39	41.1	123	US-10-952-535A-16	Sequence 16, Appl
25	39	41.1	155	US-10-952-535A-17	Sequence 17, Appl

26	39	41.1	353	US-10-055-877-30	Sequence 30, Appl
27	39	41.1	353	US-11-147-606-2	Sequence 2, Appl
28	39	41.1	408	US-11-055-822-910	Sequence 910, Appl
29	39	41.1	502	US-11-111-664-6	Sequence 6, Appl
30	39	41.1	512	US-11-111-664-8	Sequence 8, Appl
31	39	41.1	753	US-10-873-528-63	Sequence 63, Appl
32	38	40.0	38	US-11-121-612-369	Sequence 369, Appl
33	38	40.0	150	US-09-878-360A-708	Sequence 708, Appl
34	38	40.0	384	US-10-858-730-216	Sequence 216, Appl
35	38	40.0	428	US-11-138-882-8	Sequence 8, Appl
36	38	40.0	428	US-11-138-882-8	Sequence 8, Appl
37	38	40.0	483	US-11-024-959-504	Sequence 504, Appl
38	38	40.0	823	US-11-166-892-4	Sequence 4, Appl
39	37.5	39.5	13	US-11-181-148-15	Sequence 15, Appl
40	37.5	39.5	515	US-11-052-554A-255	Sequence 255, Appl
41	37	38.9	196	US-10-793-626-3070	Sequence 3070, Ap
42	37	38.9	199	US-10-793-626-186	Sequence 186, Appl
43	37	38.9	285	US-10-131-826A-448	Sequence 448, Appl
44	37	38.9	381	US-11-185-859-6	Sequence 6, Appl
45	37	38.9	431	US-11-150-533-43	Sequence 43, Appl
46	37	38.9	432	US-11-150-533-3	Sequence 3, Appl
47	37	38.9	575	US-11-150-533-5	Sequence 5, Appl
48	37	38.9	667	US-11-150-533-24	Sequence 24, Appl
49	37	38.9	668	US-10-995-561-619	Sequence 619, Appl
50	37	38.9	675	US-11-150-533-12	Sequence 12, Appl
51	37	38.9	688	US-11-150-533-10	Sequence 10, Appl
52	37	38.9	691	US-10-995-561-617	Sequence 617, Appl
53	37	38.9	692	US-11-150-533-2	Sequence 2, Appl
54	37	38.9	700	US-11-130-206-4	Sequence 4, Appl
55	37	38.9	700	US-11-130-206-2	Sequence 2, Appl
56	37	38.9	705	US-10-063-703-162	Sequence 162, Appl
57	37	38.9	705	US-11-102-240-162	Sequence 162, Appl
58	37	38.9	708	US-11-150-533-11	Sequence 11, Appl
59	37	38.9	742	US-11-150-533-65	Sequence 65, Appl
60	37	38.9	742	US-10-995-561-615	Sequence 615, Appl
61	37	38.9	742	US-10-995-561-618	Sequence 618, Appl
62	37	38.9	742	US-11-169-041-184	Sequence 184, Appl
63	37	38.9	765	US-10-131-826A-28	Sequence 28, Appl
64	37	38.9	832	US-11-098-686-10182	Sequence 10182, A
65	37	38.9	1236	US-11-080-991-68	Sequence 68, Appl
66	36	37.9	16	US-10-968-790-12	Sequence 12, Appl
67	36	37.9	225	US-11-125-295-5	Sequence 5, Appl
68	36	37.9	236	US-11-125-295-7	Sequence 7, Appl
69	36	37.9	354	US-10-485-517-179	Sequence 179, Appl
70	36	37.9	354	US-10-485-517-366	Sequence 366, Appl
71	36	37.9	356	US-10-055-877-184	Sequence 184, Appl
72	36	37.9	396	US-11-125-295-11	Sequence 11, Appl
73	36	37.9	407	US-11-125-295-9	Sequence 9, Appl
74	36	37.9	477	US-11-089-551A-34	Sequence 34, Appl
75	36	37.9	504	US-11-186-541-3	Sequence 3, Appl
76	36	37.9	507	US-10-467-657-630	Sequence 630, Appl
77	36	37.9	620	US-10-793-626-606	Sequence 606, Appl
78	36	37.9	621	US-10-793-626-2614	Sequence 2614, Ap
79	36	37.9	641	US-11-175-690-211	Sequence 211, Appl
80	36	37.9	641	US-11-175-690-230	Sequence 230, Appl
81	36	37.9	647	US-11-175-690-212	Sequence 212, Appl
82	36	37.9	653	US-11-175-690-283	Sequence 283, Appl
83	36	37.9	670	US-11-175-690-283	Sequence 283, Appl
84	36	37.9	673	US-11-175-690-201	Sequence 201, Appl
85	36	37.9	673	US-11-175-690-217	Sequence 217, Appl
86	36	37.9	673	US-11-175-690-231	Sequence 231, Appl
87	36	37.9	1115	US-10-922-232B-60	Sequence 60, Appl
88	36	37.9	6738	US-10-922-232B-15	Sequence 15, Appl
89	35.5	37.4	236	US-11-091-100-15	Sequence 15, Appl
90	35	36.8	52	US-10-895-064-2020	Sequence 2020, Ap
91	35	36.8	75	US-10-986-501-139	Sequence 139, Appl
92	35	36.8	166	US-11-176-830-1005	Sequence 1005, Ap
93	35	36.8	241	US-11-793-626-2574	Sequence 2574, Ap
94	35	36.8	269	US-11-015-546A-10	Sequence 10, Appl
95	35	36.8	308	US-11-052-554A-251	Sequence 251, Appl
96	35	36.8	311	US-10-055-877-181	Sequence 181, Appl
97	35	36.8	312	US-11-156-084-278	Sequence 278, Appl
98	35	36.8	315	US-10-524-647-25	Sequence 25, Appl

99	35	36.8	329	6	US-10-67-657-2240	Sequence 2240, Ap
100	35	36.8	329	6	US-10-524-647-2	Sequence 2, Appli

## ALIGNMENTS

```

RESULT 1
US-11-098-686-10847 1
Sequence 10847, Application US/11098686
Publication No. US20060024696A1
GENERAL INFORMATION:
APPLICANT: Kapur, Vivek and Gebhart, Connie J.
TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
FILE REFERENCE: 09531-128001
CURRENT APPLICATION NUMBER: US/11/098,686
CURRENT FILING DATE: 2005-04-04
PRIOR APPLICATION NUMBER: PCT/US03/31318
PRIOR FILING DATE: 2003-10-01
PRIOR APPLICATION NUMBER: US 60/416,335
PRIOR FILING DATE: 2002-10-04
NUMBER OF SEQ ID NOS: 11433
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10847
LENGTH: 248
TYPE: PRT
ORGANISM: Lawsonia intracellularis
US-11-098-686-10847

```

Query Match	50.0%;	Score 47.5;	DB 7;	Length 248;
Best Local Similarity	50.0%;	Pred. No. 2.1;		
Matches	8;	Conservative	4;	Mismatches 3;
				Indels 1;
				Gaps 1;

**OY**                2 LIHNV-HKEHAHAN 16  
| : | : | : ||  
**Db**              204 LHEIAHLKEHNHSN 219

```

RESULT 2
US-11-108-185-51
Sequence 51, Application US/11108185
Publication No. US20050262591A1
GENERAL INFORMATION:
APPLICANT: DeBonte, Lorin R.
APPLICANT: Fan, Zhegong
MISCELLANEOUS: Miao, Guo-Hua
TITLE OF INVENTION: FATTY ACID DESATURASES AND MUTANT SEQUENCES THEREOF
FILE REFERENCE: 07148-063003
CURRENT APPLICATION NUMBER: US/11/108,185
CURRENT FILING DATE: 2005-04-18
PRIOR APPLICATION NUMBER: US/09/771,904
PRIOR FILING DATE: 2001-01-29
PRIOR APPLICATION NUMBER: US 08/874,109
PRIOR FILING DATE: 1997-06-12
NUMBER OF SEQ ID NOS: 70
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 51
LENGTH: 29
TYPE: PRT
ORGANISM: Glycine max
US-11-108-185-51

```

Query Match	45.3%	Score 43;	DB 7;	Length 29;
Best Local Similarity	46.2%	Pred. No. 0.89;		
Matches	6;	Conservative	1;	Mismatches 6;
				Indels 0;
				Gaps 0;

QY	4	HNVHKEEHAHAHN	16
		:	
Db	12	HRTHHQHGHGAEN	24

### RESULT 3

```

US-11-015-546A-12
: Sequence 12, Application US/11/015546A
: Publication No. US20050250126A1
: GENERAL INFORMATION:
: APPLICANT: GAO, ZEREN
: APPLICANT: SHEPPARD, PAUL O.
: APPLICANT: FOX, BRIAN A.
: APPLICANT: HOLLOWAY, JAMES L.
: APPLICANT: JASPERS, STEPHEN R.
: APPLICANT: APPELBY, MARK
: TITLE OF INVENTION: Ztnf13, A TUMOR NECROSIS FACTOR
: RILE REFERENCE: 03-24
: CURRENT APPLICATION NUMBER: US/11/015,546A
: CURRENT FILING DATE: 2004-12-16
: PRIOR APPLICATION NUMBER: 60/530,185
: PRIOR FILING DATE: 2003-12-16
: NUMBER OF SEQ ID NOS: 39
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 12
: LENGTH: 274
: TYPE: PRT
: ORGANISM: Homo sapiens
US-11-015-546A-12

```

Query Match	44.2%	Score 42;	DB 7;	Length 274;
Best Local Similarity	50.0%	Pred. No. 15;		
Matches	7;	Conservative	3;	Mismatches 4;
				Indels 0;
				Gaps 0;

```
QY      2 LHNVKEEHAHAH 15
          |::: |||||
Db      101 LHHGPAQPHAAH 114
```

```

RESULT 4
US-11-116-939-14
? Sequence 14, Application US/11116939
? Publication NO. US2005026595A1
? GENERAL INFORMATION:
? APPLICANT: Stephen Tomlinson
? APPLICANT: Richard J. Quigg
? TITLE OF INVENTION: TISSUE TARGETED COMPLEMENT MODULATORS
? FILE REFERENCE: 1913.0115U2
? CURRENT APPLICATION NUMBER: US/11/116,939
? PRIORITY FILING DATE: 2005-04-28
? PRIOR APPLICATION NUMBER: 60/565,907
? PRIORITY FILING DATE: 2004-04-28
? NUMBER OF SEQ ID NOS: 27
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 14
? LENGTH: 481
? TYPE: PR1
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence; note=synthetic
? OTHER INFORMATION: construct
US-11-116-939-14

```

Query Match	44.2%	Score 42;	DB 7;	Length 481;
Best Local Similarity	53.8%	Pred. No. 28;		
Matches	7;	Conservative	3;	Mismatches 3; Indels 0; Gaps 0;

```
QY      2 LIHNVHKEHAHA 14
          | : : | | | |
Db      456 LMQQLHRELHAHA 468
```

RESULT 5  
US-11-242-243-2  
; Sequence 2, Application US/11242243  
; Publication No. US20060024605A1  
; GENERAL INFORMATION:  
; APPLICANT: Sheppard, Paul O.  
; APPLICANT: Baird, Nand



```
; APPLICANT: Deisher, Theresa A.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: DISINTEGRIN HOMOLOG
; FILE REFERENCE: 98-29
; CURRENT APPLICATION NUMBER: US/11/242,243
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/09/351,414
; PRIOR FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 696
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-242-243-2

Query Match          44.2%; Score 42; DB 7; Length 696;
Best Local Similarity 37.5%; Pred. No. 42;
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      1 HLHNHVKKEHAAH 16
Db      186 HKTYKGRSSHAHTNN 201

RESULT 6
US-10-512-109-27
; Sequence 27, Application US/10512109
; Publication No. US20050255546A1
; GENERAL INFORMATION:
; APPLICANT: KIRIN BEER KABUSHIKI KAISHA
; TITLE OF INVENTION: POLYPEPTIDE HAVING AN ACTIVITY TO SUPPORT PROLIFERATION OR SURVIV
; TITLE OF INVENTION: OF HEMATOPOIETIC STEM CELL OR HEMATOPOIETIC PROGENITOR CELL, AND
; FILE REFERENCE: 905M01OP1572
; CURRENT APPLICATION NUMBER: US/10/512,109
; CURRENT FILING DATE: 2004-10-21
; PRIOR APPLICATION NUMBER: US 60/376,001
; PRIOR FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-512-109-27

Query Match          44.2%; Score 42; DB 6; Length 829;
Best Local Similarity 37.5%; Pred. No. 51;
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      1 HLHNHVKKEHAAH 16
Db      306 HKTYKGRSSHAHTNN 321

RESULT 7
US-10-512-109-29
; Sequence 29, Application US/10512109
; Publication No. US20050255546A1
; GENERAL INFORMATION:
; APPLICANT: KIRIN BEER KABUSHIKI KAISHA
; TITLE OF INVENTION: POLYPEPTIDE HAVING AN ACTIVITY TO SUPPORT PROLIFERATION OR SURVIV
; TITLE OF INVENTION: OF HEMATOPOIETIC STEM CELL OR HEMATOPOIETIC PROGENITOR CELL, AND
; FILE REFERENCE: 905M01OP1572
; CURRENT APPLICATION NUMBER: US/10/512,109
; CURRENT FILING DATE: 2004-10-21
; PRIOR APPLICATION NUMBER: US 60/376,001
; PRIOR FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
```

```
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-512-109-29

Query Match          44.2%; Score 42; DB 6; Length 832;
Best Local Similarity 37.5%; Pred. No. 51;
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      1 HLHNHVKKEHAAH 16
Db      309 HKTYKGRSSHAHTNN 324

RESULT 8
US-11-054-281-104
; Sequence 104, Application US/11054281
; Publication No. US20060013813A1
; GENERAL INFORMATION:
; APPLICANT: Mezes et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-240CIP
; CURRENT APPLICATION NUMBER: US/11/054,281
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: 60/261,014
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,018
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/318,410
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/261,013
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,026
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,029
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/313,170
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 10/044,564
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 104
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-054-281-104

Query Match          43.7%; Score 41.5; DB 7; Length 382;
Best Local Similarity 35.3%; Pred. No. 26;
Matches 6; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY      1 HLI-HNVHKEHAAH 16
Db      190 HLIHHHHQVHPHSHS 206

RESULT 9
US-11-054-281-103
; Sequence 103, Application US/11054281
; Publication No. US20060013813A1
; GENERAL INFORMATION:
; APPLICANT: Mezes et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-240CIP
; CURRENT APPLICATION NUMBER: US/11/054,281
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: 60/261,014
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,018
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/318,410
; PRIOR FILING DATE: 2001-09-10
```

Query Match	43.7%	Score 41.5,	DB 6,	Length 755,
Best Local Similarity	35.3%	Pred. No. 54,		
Matches 6;	Conservative 6;	Mismatches 4;	Indels 1;	Gaps 1

Db 563 HILHHHQNHHPHSHS 579

RESULT 13

US-11-054-281-26  
 ; Sequence 26, Application US/11054281  
 ; Publication No. US20060013813A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mezes et al.  
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
 ; FILE REFERENCE: 21402-240CIP  
 ; CURRENT APPLICATION NUMBER: US/11/054,281  
 ; CURRENT FILING DATE: 2005-02-08  
 ; PRIOR APPLICATION NUMBER: 60/261,014  
 ; PRIOR FILING DATE: 2001-01-11  
 ; PRIOR APPLICATION NUMBER: 60/261,018  
 ; PRIOR FILING DATE: 2001-01-11  
 ; PRIOR APPLICATION NUMBER: 60/318,410  
 ; PRIOR FILING DATE: 2001-09-10  
 ; PRIOR APPLICATION NUMBER: 60/261,013  
 ; PRIOR FILING DATE: 2001-01-11  
 ; PRIOR APPLICATION NUMBER: 60/261,026  
 ; PRIOR FILING DATE: 2001-01-11  
 ; PRIOR APPLICATION NUMBER: 60/261,029  
 ; PRIOR FILING DATE: 2001-01-11  
 ; PRIOR APPLICATION NUMBER: 60/313,170  
 ; PRIOR FILING DATE: 2001-08-17  
 ; PRIOR APPLICATION NUMBER: 10/044,564  
 ; PRIOR FILING DATE: 2002-01-11  
 ; NUMBER OF SEQ ID NOS: 324  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 26  
 ; LENGTH: 755  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-11-054-281-26

Query Match 43.7%; Score 41.5; DB 7; Length 755;  
 Best Local Similarity 35.3%; Pred. No. 54;  
 Matches 6; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 1 HLI-HNVKKEHNAH 16  
 Db 563 HILHHHQNHHPHSHS 579

RESULT 14

US-11-054-281-102  
 ; Sequence 102, Application US/11054281  
 ; Publication No. US20060013813A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mezes et al.  
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
 ; FILE REFERENCE: 21402-240CIP  
 ; CURRENT APPLICATION NUMBER: US/11/054,281  
 ; CURRENT FILING DATE: 2005-02-08  
 ; PRIOR APPLICATION NUMBER: 60/261,014  
 ; PRIOR FILING DATE: 2001-01-11  
 ; PRIOR APPLICATION NUMBER: 60/261,018  
 ; PRIOR FILING DATE: 2001-01-11  
 ; PRIOR APPLICATION NUMBER: 60/318,410  
 ; PRIOR FILING DATE: 2001-09-10  
 ; PRIOR APPLICATION NUMBER: 60/261,013  
 ; PRIOR FILING DATE: 2001-01-11  
 ; PRIOR APPLICATION NUMBER: 60/261,026  
 ; PRIOR FILING DATE: 2001-01-11  
 ; PRIOR APPLICATION NUMBER: 60/261,029  
 ; PRIOR FILING DATE: 2001-01-11  
 ; PRIOR APPLICATION NUMBER: 60/313,170  
 ; PRIOR FILING DATE: 2001-08-17  
 ; PRIOR APPLICATION NUMBER: 10/044,564  
 ; PRIOR FILING DATE: 2002-01-11  
 ; NUMBER OF SEQ ID NOS: 324

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 102  
 ; LENGTH: 755  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-11-054-281-102

Query Match 43.7%; Score 41.5; DB 7; Length 755;  
 Best Local Similarity 35.3%; Pred. No. 54;  
 Matches 6; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 1 HLI-HNVKKEHNAH 16  
 Db 563 HILHHHQNHHPHSHS 579

RESULT 15

US-10-968-790-27  
 ; Sequence 27, Application US/10968790  
 ; Publication No. US20050287186A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ellis-Behnke, et al.  
 ; TITLE OF INVENTION: Self-Assembling Peptides For Repair and Regeneration of Neural  
 ; FILE REFERENCE: 0492611-0590  
 ; CURRENT APPLICATION NUMBER: US/10/968,790  
 ; CURRENT FILING DATE: 2004-10-18  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 27  
 ; LENGTH: 16  
 ; TYPE: PRT  
 ; ORGANISM: Artificial  
 ; FEATURE:  
 ; OTHER INFORMATION: Self-Assembling Peptides  
 ; US-10-968-790-27

Query Match 43.2%; Score 41; DB 6; Length 16;  
 Best Local Similarity 46.7%; Pred. No. 0.91;  
 Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 HLI-HNVKKEHNAH 15  
 Db 1 HEHEHKKHEHKKH 15

RESULT 16

US-11-124-368A-336  
 ; Sequence 336, Application US/11124368A  
 ; Publication No. US20050287559A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Michele Cargili  
 ; APPLICANT: James J. Devlin  
 ; TITLE OF INVENTION: Genetic Polymorphisms Associated with  
 ; FILE REFERENCE: C1001524  
 ; CURRENT APPLICATION NUMBER: US/11/124,368A  
 ; CURRENT FILING DATE: 2005-05-09  
 ; PRIOR APPLICATION NUMBER: US 60/568,845  
 ; PRIOR FILING DATE: 2004-05-07  
 ; PRIOR APPLICATION NUMBER: US 60/525,936  
 ; PRIOR FILING DATE: 2004-11-09  
 ; NUMBER OF SEQ ID NOS: 21112  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 336  
 ; LENGTH: 469  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-11-124-368A-336

Query Match 42.1%; Score 40; DB 7; Length 469;  
 Best Local Similarity 42.9%; Pred. No. 53;

Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 LHNHVKKEHNAH 15

Db 101 LVHRGHGHDEHSH 114

RESULT 17

US-11-124-368A-337  
Sequence 337, Application US/11124368A  
Publication No. US20050287559A1  
GENERAL INFORMATION:  
APPLICANT: Michele Cargill  
APPLICANT: James J. Devlin  
APPLICANT: May Luke  
TITLE OF INVENTION: Genetic Polymorphisms Associated with  
TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof  
FILE REFERENCE: CL001524  
CURRENT APPLICATION NUMBER: US/11/124,368A  
PRIOR FILING DATE: 2005-05-09  
PRIOR APPLICATION NUMBER: US 60/568,845  
PRIOR FILING DATE: 2004-05-07  
PRIOR APPLICATION NUMBER: US 60/625,936  
NUMBER OF SEQ ID NOS: 21112  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 337  
LENGTH: 469  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-11-124-368A-337

Query Match 42.1%; Score 40; DB 7; Length 469;

Best Local Similarity 42.9%; Pred. No. 53;

Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 LHNHVKKEHNAH 15

Db 101 LVHRGHGHDEHSH 114

RESULT 18

US-11-185-111-41  
Sequence 41, Application US/11185111  
Publication No. US20050255519A1  
GENERAL INFORMATION:  
APPLICANT: BILLING-MEDEL, PATRICIA  
APPLICANT: COHEN, MAURICE  
APPLICANT: COLPITS, TRACEY L.  
APPLICANT: FRIEDMAN, PAULA N.  
APPLICANT: GORDON, JULIAN  
APPLICANT: GRANADOS, EDWARD N.  
APPLICANT: HODGES, STEVEN C.  
APPLICANT: KRATOCHVIL, JON D.  
APPLICANT: KRAUS, MICHAEL R.  
APPLICANT: ROBERTS-RAPP, LISA  
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL  
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/11/185,111

FILING DATE: 20-JULY-2005

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/841,894

FILING DATE: 25-Apr-2001

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/071,710

FILING DATE: <Unknown>

APPLICATION NUMBER: 08/850,713

FILING DATE: 02-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Becker, Cheryl L.

REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6063.US.P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729

TELEFAX: 847/938-2623

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 41:

SEQUENCE CHARACTERISTICS:

LENGTH: 21 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: None

US-11-185-111-41

Query Match 41.1%; Score 39; DB 7; Length 21;

Best Local Similarity 54.5%; Pred. No. 2.4;

Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 5 NVHKEHNAH 15

Db 11 NVHTEHHHHH 21

RESULT 19

US-11-148-108-51  
Sequence 51, Application US/11148108  
Publication No. US2005026477A1  
GENERAL INFORMATION:  
APPLICANT: BILLING-MEDEL, PATRICIA A.  
APPLICANT: COHEN, MAURICE  
APPLICANT: COLPITS, TRACEY L.  
APPLICANT: FRIEDMAN, PAULA N.  
APPLICANT: HAYDEN, MARK  
APPLICANT: KRAUS, MICHAEL R.  
APPLICANT: ROBERTS-RAPP, LISA  
APPLICANT: RUSSELL, JOHN C.  
APPLICANT: STROUPE, STEPHEN D.  
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL  
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE GASTROINTESTINAL  
TITLE OF INVENTION: TRACT  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/11/148,108  
FILING DATE: 08-JUN-2005  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/049,696

FILING DATE: 27-MAR-1998  
APPLICATION NUMBER: 06/829,754  
FILING DATE: 31-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Becker, Cheryl L.  
REGISTRATION NUMBER: 35,441  
REFERENCE/DOCKET NUMBER: 6067.US.P1  
TELEPHONE: 847/935-1729  
TELEFAX: 847/938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 51:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-11-148-106-51

Query Match 41.1%; Score 39; DB 7; Length 21;  
Best Local Similarity 54.5%; Pred. No. 2.4;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 NVHKEEHAAH 15  
DB 11 NMATEHHHHH 21

RESULT 20  
US-11-094-071-31  
Sequence 31, Application US/11094071  
Publication No. US20050282187A1  
GENERAL INFORMATION:  
APPLICANT: BILLING-MEDEL, PATRICIA  
APPLICANT: COHEN, MAURICE  
APPLICANT: COLETT, TRACEY L.  
APPLICANT: FRIEDMAN, PAULA N.  
APPLICANT: GRANADOS, EDWARD N.  
APPLICANT: HODGES, STEVEN C.  
APPLICANT: KLAS, MICHAEL R.  
APPLICANT: KRATOCHVIL, JON D.  
APPLICANT: ROBERTS-RAP, LISA  
APPLICANT: RUSSELL, JOHN C.  
APPLICANT: STROUB, STEPHEN D.  
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL  
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE URINARY TRACT  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/11/094,071  
FILING DATE: 30-Mar-2005  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/079,874  
FILING DATE: 15-May-1998  
APPLICATION NUMBER: 08/856,652  
FILING DATE: 15-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Becker, Cheryl L.  
REGISTRATION NUMBER: 35,441  
REFERENCE/DOCKET NUMBER: 6106.US.P1  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729  
TELEFAX: 847/938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: None  
US-11-094-071-31

Query Match 41.1%; Score 39; DB 7; Length 21;  
Best Local Similarity 54.5%; Pred. No. 2.4;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 NVHKEEHAAH 15  
DB 11 NMATEHHHHH 21

RESULT 21  
US-11-108-185-49  
Sequence 49, Application US/11108185  
Publication No. US20050262591A1  
GENERAL INFORMATION:  
APPLICANT: DeBonte, Lorin R.  
APPLICANT: Fan, Zhongong  
APPLICANT: Miao, Guo-Hua  
TITLE OF INVENTION: FATTY ACID DESATURASES AND MUTANT SEQUENCES THEREOF  
FILE REFERENCE: 07148-063003  
CURRENT APPLICATION NUMBER: US/11/108,185  
CURRENT FILING DATE: 2005-04-18  
PRIOR APPLICATION NUMBER: US/09/771,904  
PRIOR FILING DATE: 2001-01-29  
PRIOR APPLICATION NUMBER: US 08/874,109  
PRIOR FILING DATE: 1997-06-12  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 49  
LENGTH: 29  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-11-108-185-49

Query Match 41.1%; Score 39; DB 7; Length 29;  
Best Local Similarity 38.5%; Pred. No. 3.4;  
Matches 5; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 4 NVHKEEHAAH 16  
DB 12 HRTTHQHGHVH 24

RESULT 22  
US-11-108-185-52  
Sequence 52, Application US/11108185  
Publication No. US20050262591A1  
GENERAL INFORMATION:  
APPLICANT: DeBonte, Lorin R.  
APPLICANT: Fan, Zhongong  
APPLICANT: Miao, Guo-Hua  
TITLE OF INVENTION: FATTY ACID DESATURASES AND MUTANT SEQUENCES THEREOF  
FILE REFERENCE: 07148-063003  
CURRENT APPLICATION NUMBER: US/11/108,185  
CURRENT FILING DATE: 2005-04-18  
PRIOR APPLICATION NUMBER: US/09/771,904  
PRIOR FILING DATE: 2001-01-29  
PRIOR APPLICATION NUMBER: US 08/874,109  
PRIOR FILING DATE: 1997-06-12  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 52

/ LENGTH: 29  
/ TYPE: PRT  
/ ORGANISM: Arabidopsis thaliana  
US-11-108-185-52

Query Match 41.1%; Score 39; DB 7; Length 29;  
Best Local Similarity 38.5%; Pred. No. 3.4;  
Matches 5; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 4 NVHKEEHAH 16  
|:|:|:|:|  
DB 12 HRTTHQNHGIVEN 24

RESULT 23  
US-10-952-535A-15  
/ Sequence 15, Application US/10952535A  
/ Publication No. US20050255113A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Huston, James S.  
/ APPLICANT: Messer, Anne  
/ APPLICANT: Lecerf, Jean-Michel  
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING POLYPEPTIDE  
/ FILE REFERENCE: INR-004CP  
/ CURRENT APPLICATION NUMBER: US/10/952,535A  
/ PRIOR FILING DATE: 2004-09-27  
/ PRIOR APPLICATION NUMBER: 60/146,047  
/ NUMBER OF SEQ ID NOS: 45  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 15  
/ LENGTH: 98  
/ TYPE: PRT  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-952-535A-15

Query Match 41.1%; Score 39; DB 6; Length 98;  
Best Local Similarity 54.5%; Pred. No. 13;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 NVHKEEHAH 15  
|:|:|:|:|  
DB 88 NMHTTEHHHH 98

RESULT 24  
US-10-952-535A-16  
/ Sequence 16, Application US/10952535A  
/ Publication No. US20050255113A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Huston, James S.  
/ APPLICANT: Messer, Anne  
/ APPLICANT: Lecerf, Jean-Michel  
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING POLYPEPTIDE  
/ FILE REFERENCE: INR-004CP  
/ CURRENT APPLICATION NUMBER: US/10/952,535A  
/ PRIOR FILING DATE: 2004-09-27  
/ PRIOR APPLICATION NUMBER: 60/146,047  
/ NUMBER OF SEQ ID NOS: 45  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 16  
/ LENGTH: 123  
/ TYPE: PRT  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
/ OTHER INFORMATION: Construct

US-10-952-535A-16

Query Match 41.1%; Score 39; DB 6; Length 123;  
Best Local Similarity 54.5%; Pred. No. 17;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 NVHKEEHAH 15  
|:|:|:|:|  
DB 113 NMHTTEHHHH 123

RESULT 25  
US-10-952-535A-17  
/ Sequence 17, Application US/10952535A  
/ Publication No. US20050255113A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Huston, James S.  
/ APPLICANT: Messer, Anne  
/ APPLICANT: Lecerf, Jean-Michel  
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING POLYPEPTIDE  
/ FILE REFERENCE: INR-004CP  
/ CURRENT APPLICATION NUMBER: US/10/952,535A  
/ PRIOR FILING DATE: 2004-09-27  
/ PRIOR APPLICATION NUMBER: 60/146,047  
/ NUMBER OF SEQ ID NOS: 45  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 17  
/ LENGTH: 155  
/ TYPE: PRT  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-952-535A-17

Query Match 41.1%; Score 39; DB 6; Length 155;  
Best Local Similarity 54.5%; Pred. No. 22;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 NVHKEEHAH 15  
|:|:~|:~|:~|  
DB 145 NMHTTEHHHH 155

RESULT 26  
US-10-055-877-30  
/ Sequence 30, Application US/10055877  
/ Publication No. US20050288241A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Decristofaro, Marc  
/ APPLICANT: Padigaru, Muraidhara  
/ APPLICANT: Miller, Charles  
/ APPLICANT: Tchervnev, Velizar  
/ APPLICANT: Zhong, Mei  
/ APPLICANT: Anderson, David  
/ APPLICANT: Ballinger, Robert  
/ APPLICANT: Gerlach, Valerie  
/ APPLICANT: Spytek, Kimberly  
/ APPLICANT: Ratalli, Luca  
/ APPLICANT: Kehuda, Ramesh  
/ APPLICANT: Guo, Xiaojia  
/ APPLICANT: Zernussen, Bryan  
/ APPLICANT: Andrew, David  
/ APPLICANT: Mezes, Peter  
/ APPLICANT: Patnurejan, Meera  
/ APPLICANT: Burgess, Catherine  
/ APPLICANT: Eisen, Andrew  
/ APPLICANT: Wolenc, Adam  
/ APPLICANT: Baumgartner, Jason  
/ APPLICANT: Shinkets, Richard  
/ APPLICANT: Gusev, Vladimir

```
; APPLICANT: Vernet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-055-877-30
```

```
Query Match          41.1%; Score 39; DB 6; Length 353;
Best Local Similarity 37.5%; Pred. No. 55;
Matches 6; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
```

```
QY      1 HLHNVKKEHAAH 16
Db      190 HAAHHHAAHHHHH 205
```

```
RESULT 27
US-11-147-606-2
; Sequence 2, Application US/11147606
; Publication No. US20050278795A1
; GENERAL INFORMATION:
; APPLICANT: Ntambi, James M.
; APPLICANT: Miyazaki, Makoto
; TITLE OF INVENTION: Stearoyl-CoA desaturase 4 gene
; FILE REFERENCE: 960296,00197
; CURRENT APPLICATION NUMBER: US/11/147,606
; CURRENT FILING DATE: 2005-06-08
; PRIOR APPLICATION NUMBER: 60/578,234
; PRIOR FILING DATE: 2004-06-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-147-606-2
```

```
Query Match          41.1%; Score 39; DB 7; Length 353;
Best Local Similarity 43.8%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
```

```
QY      1 HLHNVKKEHAAH 16
Db      151 HRAHKRFETHADPH 166
```

```
RESULT 28
US-11-055-822-910
; Sequence 910, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krogger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberkauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: BGI-121CPCN
; CURRENT APPLICATION NUMBER: US/11/055,822
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 09/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 910
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-910
```

```
Query Match          41.1%; Score 39; DB 7; Length 408;
Best Local Similarity 53.8%; Pred. No. 64;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      3 IHNVKKEHAAH 15
Db      41 VHAVDREHAAH 53
```

```
RESULT 29
US-11-111-664-6
; Sequence 6, Application US/11111664
; Publication No. US20060014687A1
; GENERAL INFORMATION:
; APPLICANT: ENOBIA PHARMA INC.
; APPLICANT: Crine, Philippe
; APPLICANT: Loisel, Thomas
; APPLICANT: Lemire, Isabelle
; APPLICANT: Boileau, Guy
; TITLE OF INVENTION: BONE DELIVERY CONJUGATES AND METHOD OF USING SAME TO TARGET
; FILE REFERENCE: 2006078-0003
; CURRENT APPLICATION NUMBER: US/11/111,664
; CURRENT FILING DATE: 2005-04-21
```

```
/ PRIOR APPLICATION NUMBER: US 60/563,828
/ PRIOR FILING DATE: 2004-04-21
/ PRIOR APPLICATION NUMBER: US 60/614,984
/ PRIOR FILING DATE: 2004-10-04
/ PRIOR APPLICATION NUMBER: US 60/590,347
/ PRIOR FILING DATE: 2004-07-23
/ NUMBER OF SEQ ID NOS: 44
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 6
/ LENGTH: 502
/ TYPE: PRT
/ ORGANISM: Artificial
/ FEATURE:
/ OTHER INFORMATION: Soluble alkaline phosphatase
US-11-111-664-6
```

```
Query Match      41.1%; Score 39; DB 7; Length 502;
Best Local Similarity 45.5%; Pred. No. 80;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 HLHNVKKEH 11
      ||| ||| |||
Db      469 HLHGVHEQNY 479
```

```
RESULT 30
US-11-111-664-8
/ Sequence 8, Application US/11111664
/ Publication No. US20060014687A1
/ GENERAL INFORMATION:
/ APPLICANT: ENOBIA PHARMA INC.
/ APPLICANT: Crine, Philippe
/ APPLICANT: Loisel, Thomas
/ APPLICANT: Lemire, Isabelle
/ APPLICANT: Boileau, Guy
/ TITLE OF INVENTION: BONE DELIVERY CONJUGATES AND METHOD OF USING SAME TO TARGET
/ FILE REFERENCE: 2006078-0003
/ CURRENT APPLICATION NUMBER: US/11/111,664
/ CURRENT FILING DATE: 2005-04-21
/ PRIOR APPLICATION NUMBER: US 60/563,828
/ PRIOR FILING DATE: 2004-04-21
/ PRIOR APPLICATION NUMBER: US 60/614,984
/ PRIOR FILING DATE: 2004-10-04
/ PRIOR APPLICATION NUMBER: US 60/590,347
/ PRIOR FILING DATE: 2004-07-23
/ NUMBER OF SEQ ID NOS: 44
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 8
/ LENGTH: 512
/ TYPE: PRT
/ ORGANISM: Artificial
/ FEATURE:
/ OTHER INFORMATION: Soluble alkaline phosphatase conjugate
US-11-111-664-8
```

```
Query Match      41.1%; Score 39; DB 7; Length 512;
Best Local Similarity 45.5%; Pred. No. 82;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 HLHNVKKEH 11
      ||| ||| |||
Db      469 HLHGVHEQNY 479
```

```
RESULT 31
US-10-873-528-63
/ Sequence 63, Application US/10873528
/ Publication No. US20050276814A1
/ GENERAL INFORMATION:
/ APPLICANT: Microdial Technics Limited
/ APPLICANT: Gilbert, Christophe
/ APPLICANT: Hansbro, Philip M
```

```
/ TITLE OF INVENTION: Proteins
/ FILE REFERENCE: PWC/P21129WO
/ CURRENT APPLICATION NUMBER: US/10/873,528
/ CURRENT FILING DATE: 2004-06-23
/ PRIOR APPLICATION NUMBER: US/09/769,787
/ PRIOR FILING DATE: 2001-01-26
/ PRIOR APPLICATION NUMBER: GB 9816337.1
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: US 60/125164
/ PRIOR FILING DATE: 1999-03-19
/ NUMBER OF SEQ ID NOS: 388
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 63
/ LENGTH: 763
/ TYPE: PRT
/ ORGANISM: Streptococcus pneumoniae
US-10-873-528-63
```

```
Query Match      41.1%; Score 39; DB 6; Length 763;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      3 HNVKKEHA 12
      ||| ||| |||
Db      248 IDLYKKEHA 257
```

```
RESULT 32
US-11-121-612-369
/ Sequence 369, Application US/11121612
/ Publication No. US20060025339A1
/ GENERAL INFORMATION:
/ APPLICANT: The Procter & Gamble Company
/ APPLICANT: Isfort, Robert J
/ APPLICANT: Mazur, Wieslaw A
/ TITLE OF INVENTION: Corticotropin Releasing Factor 2 Receptor Agonists
/ FILE REFERENCE: 8847MD
/ CURRENT APPLICATION NUMBER: US/11/121,612
/ CURRENT FILING DATE: 2005-05-04
/ PRIOR APPLICATION NUMBER: US 60/349,117
/ PRIOR FILING DATE: 2002-01-16
/ PRIOR APPLICATION NUMBER: US 60/376,337
/ PRIOR FILING DATE: 2002-04-29
/ PRIOR APPLICATION NUMBER: US 60/388,895
/ PRIOR FILING DATE: 2002-06-14
/ PRIOR APPLICATION NUMBER: US 60/411,988
/ PRIOR FILING DATE: 2002-09-19
/ PRIOR APPLICATION NUMBER: US 10/317,252
/ PRIOR FILING DATE: 2002-12-11
/ NUMBER OF SEQ ID NOS: 531
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 369
/ LENGTH: 38
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Chemically synthesized peptide sequence
/ NAME/KEY: MOD_RBS
/ LOCATION: (38)..(38)
/ OTHER INFORMATION: AMIDATION
US-11-121-612-369
```

```
Query Match      40.0%; Score 38; DB 7; Length 38;
Best Local Similarity 66.7%; Pred. No. 6.5;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      7 HKEKHAHAH 15
      ||| ||| |||
Db      21 HAHAAHAH 29
```

RESULT 33



US-09-978-360A-708  
; Sequence 708, Application US/09978360A  
; Publication No. US20060009633A9  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; APPLICANT: Clusell, Catherine  
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with signal peptides  
; FILE REFERENCE: 56.US4.CIP  
; CURRENT APPLICATION NUMBER: US/09/978,360A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: US 60/066,677  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: US 60/069,957  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: US 60/074,121  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: US 60/081,563  
; PRIOR FILING DATE: 1998-04-13  
; PRIOR APPLICATION NUMBER: US 60/096,116  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: US 60/099,273  
; PRIOR FILING DATE: -09-04  
; PRIOR APPLICATION NUMBER: US 09/191,997  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: US 09/215,435  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: PCT/IB98/02122  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: US 09/247,155  
; PRIOR FILING DATE: 1999-02-09  
; Remaining prior Application data removed - See file Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 810  
; SOFTWARE: Patent.pm  
; SEQ ID NO 708  
; LENGTH: 150  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-978-360A-708

Query Match 40.0%; Score 38; DB 5; Length 150;  
Best Local Similarity 50.0%; Pred. No. 30;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 7 HKEEHAAHN 16  
| | | | |  
Db 124 HSHNHQSHN 133

RESULT 34  
US-10-858-730-216  
; Sequence 216, Application US/10858730  
; Publication No. US20050255568A1  
; GENERAL INFORMATION:  
; APPLICANT: Bailey, Richard B.  
; APPLICANT: Blomquist, Paul  
; APPLICANT: Doten, Reed  
; APPLICANT: Driggers, Edward M.  
; APPLICANT: Madden, Kevin T.  
; APPLICANT: O'Leary, Jessica  
; APPLICANT: O'Toole, George  
; APPLICANT: Trueheart, Joshua  
; APPLICANT: Walbridge, Michael J.  
; APPLICANT: Vorgey, Peter S.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID  
; TITLE OF INVENTION: PRODUCTION  
; FILE REFERENCE: 14184-030001  
; CURRENT APPLICATION NUMBER: US/10/858,730  
; CURRENT FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/475,000  
; PRIOR FILING DATE: 2003-05-30

; PRIOR APPLICATION NUMBER: US 60/551,860  
; PRIOR FILING DATE: 2004-03-10  
; NUMBER OF SEQ ID NOS: 364  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 216  
; LENGTH: 384  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-858-730-216

Query Match 40.0%; Score 38; DB 6; Length 384;  
Best Local Similarity 42.9%; Pred. No. 84;  
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 LHHVKEEHAAHN 15  
| | | | |  
Db 347 LHHVKEETAAYGH 360

RESULT 35  
US-11-138-642-8  
; Sequence 8, Application US/11138642  
; Publication No. US20050287190A1  
; GENERAL INFORMATION:  
; APPLICANT: DESOUZA, MERVYN L.  
; APPLICANT: MESSMAN, MICHAEL A.  
; TITLE OF INVENTION: METHODS OF SUPPRESSING ENDOTOXIN EFFECTS IN ANIMAL  
; TITLE OF INVENTION: FEEDS CONTAINING E. COLI BIOMASS  
; FILE REFERENCE: 023829-0454  
; CURRENT APPLICATION NUMBER: US/11/138,642  
; CURRENT FILING DATE: 2005-05-26  
; PRIOR APPLICATION NUMBER: 60/575,272  
; PRIOR FILING DATE: 2004-05-28  
; PRIOR APPLICATION NUMBER: 60/578,092  
; PRIOR FILING DATE: 2004-06-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Patentin Ver. 3.3  
; SEQ ID NO 8  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-11-138-642-8

Query Match 40.0%; Score 38; DB 7; Length 428;  
Best Local Similarity 41.2%; Pred. No. 95;  
Matches 7; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

QY 1 HLHHVKEE--HAAHN 15  
| | | | |  
Db 197 HSHSKEDDELHAAHS 213

RESULT 36  
US-11-138-882-8  
; Sequence 8, Application US/11138882  
; Publication No. US20060008546A1  
; GENERAL INFORMATION:  
; APPLICANT: DESOUZA, MERVYN L.  
; APPLICANT: MESSMAN, MICHAEL A.  
; TITLE OF INVENTION: ORGANISMS WITH ENHANCED HISTIDINE BIOSYNTHESIS AND  
; TITLE OF INVENTION: THEIR USE IN ANIMAL FEEDS  
; FILE REFERENCE: 023829-0451  
; CURRENT APPLICATION NUMBER: US/11/138,882  
; CURRENT FILING DATE: 2005-05-26  
; PRIOR APPLICATION NUMBER: 60/575,470  
; PRIOR FILING DATE: 2004-05-28  
; PRIOR APPLICATION NUMBER: 60/578,098  
; PRIOR FILING DATE: 2004-06-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Patentin Ver. 3.3  
; SEQ ID NO 8  
; LENGTH: 428  
; TYPE: PRT

ORGANISM: Mus musculus  
US-11-138-882-8

Query Match 40.0%; Score 38; DB 7; Length 428;  
Best Local Similarity 41.2%; Pred. No. 95;  
Matches 7; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

QY 1 HHHVHKEE--HAAH 15  
DB 197 HSHSKERDELHAAH 213

RESULT 37  
US-11-024-959-504  
Sequence 504, Application US/11024959  
Publication No. US20060010516A1

GENERAL INFORMATION:  
APPLICANT: FORSTER, RICHARD L.  
APPLICANT: CONNETT, MARIE B.  
APPLICANT: EMERSON, SARAH JANE  
APPLICANT: GRIGOR, MURRAY ROBERT  
APPLICANT: HIGGINS, COLLEEN M.  
APPLICANT: LUND, STEVEN TROY  
APPLICANT: MAGUSIN, ANDREAS  
APPLICANT: KODRZYCKI, BOB  
TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS  
FILE REFERENCE: 044463-0360  
CURRENT APPLICATION NUMBER: US/11/024,959  
CURRENT FILING DATE: 2004-12-30  
PRIOR APPLICATION NUMBER: 60/533,036  
PRIOR FILING DATE: 2003-12-30  
NUMBER OF SEQ ID NOS: 782  
SOFTWARE: PatentIn version 3.3

SEQ ID NO 504  
LENGTH: 483  
TYPE: PRT  
ORGANISM: Pinus radiata  
US-11-024-959-504

Query Match 40.0%; Score 38; DB 7; Length 483;  
Best Local Similarity 46.2%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 4 HNVHKEEHAH 16  
DB 30 HNVHNSHSHYH 42

RESULT 38  
US-11-166-892-4  
Sequence 4, Application US/11166892  
Publication No. US20060014705A1

GENERAL INFORMATION:  
APPLICANT: Howitz, Konrad T  
APPLICANT: Zupkin, Robert E  
TITLE OF INVENTION: Compositions and Methods for Selectively Activating Human Sitruin  
FILE REFERENCE: BIO00002HS  
CURRENT APPLICATION NUMBER: US/11/166,892  
CURRENT FILING DATE: 2005-06-24  
PRIOR APPLICATION NUMBER: US 60/584,943  
PRIOR FILING DATE: 2004-06-30  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 4  
LENGTH: 823  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-11-166-892-4

Query Match 40.0%; Score 38; DB 7; Length 823;  
Best Local Similarity 40.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 HHHVHKEEHAH 15  
DB 508 HSHHLLHHHHRH 522

RESULT 39  
US-11-181-148-15  
Sequence 15, Application US/11181148  
Publication No. US20060019292A1

GENERAL INFORMATION:  
APPLICANT: Farmer, Alan Andrew  
TITLE OF INVENTION: SEQUENCE SPECIFIC RECOMBINASE-BASED  
METHODS FOR PRODUCING INTRON CONTAINING VECTORS AND  
TITLE OF INVENTION: COMPOSITIONS FOR USE IN PRACTICING THE SAME  
FILE REFERENCE: CLON-069CON  
CURRENT APPLICATION NUMBER: US/11/181,148  
CURRENT FILING DATE: 2005-07-13  
PRIOR APPLICATION NUMBER: 60/263,358  
PRIOR FILING DATE: 2001-01-16  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 13  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: tag sequence  
US-11-181-148-15

Query Match 39.5%; Score 37.5; DB 7; Length 13;  
Best Local Similarity 53.8%; Pred. No. 2.4;  
Matches 7; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 4 HNVHKEEHAH 16  
DB 2 HN-HNHNHNHN 13

RESULT 40  
US-11-052-554A-255  
Sequence 255, Application US/11052554A  
Publication No. US20050288866A1  
GENERAL INFORMATION:  
APPLICANT: Sachdeva, et al.  
TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
PROTEINS OF THERAPEUTIC POTENTIAL  
FILE REFERENCE: 30853/40359A  
CURRENT APPLICATION NUMBER: US/11/052,554A  
CURRENT FILING DATE: 2005-02-07  
PRIOR APPLICATION NUMBER: US 60/589,227  
PRIOR FILING DATE: 2004-07-20  
PRIOR APPLICATION NUMBER: IN 173/DEL/2004  
PRIOR FILING DATE: 2004-02-06  
NUMBER OF SEQ ID NOS: 763  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 255  
LENGTH: 515  
TYPE: PRT  
ORGANISM: Streptococcus pyogenes MGAS8232  
US-11-052-554A-255

Query Match 39.5%; Score 37.5; DB 7; Length 515;  
Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
Matches 8; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 4 HNVHKEE-HAHA 14  
DB 137 HNHSHSEGHNHA 148

Search completed: February 11, 2006, 13:20:03  
Job time : 18 secs

November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New).

Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

*This page blank (here)*

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2006, 13:15:58 / Search time 178 Seconds  
(Without alignments)  
37,558 Million cell updates/sec

Title: US-10-762-588-1

Perfect score: 95  
Sequence: 1 HLHNVKHEHMAHN 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Published Applications AA Main:\*

- 1: /cgn2\_6/pcodata/1/pubppaa/us07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/pcodata/1/pubppaa/us08\_PUBCOMB.pep:\*
- 3: /cgn2\_6/pcodata/1/pubppaa/us09\_PUBCOMB.pep:\*
- 4: /cgn2\_6/pcodata/1/pubppaa/us10A\_PUBCOMB.pep:\*
- 5: /cgn2\_6/pcodata/1/pubppaa/us10B\_PUBCOMB.pep:\*
- 6: /cgn2\_6/pcodata/1/pubppaa/us11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	100.0	16	3 US-09-858-332-2	Sequence 2, Appl1
2	95	100.0	16	4 US-10-762-588-1	Sequence 1, Appl1
3	95	100.0	20	4 US-10-748-055-1	Sequence 1, Appl1
4	95	100.0	279	3 US-09-858-332-15	Sequence 15, Appl1
5	87	91.6	19	5 US-10-634-645-14	Sequence 14, Appl1
6	53	55.8	837	6 US-11-097-143-8130	Sequence 8130, Ap
7	51	53.7	402	6 US-11-097-143-20379	Sequence 20379, A
8	51	53.7	565	4 US-10-425-115-344916	Sequence 344916,
9	52	60	4	US-10-425-115-365836	Sequence 365836,
10	49	51.6	124	5 US-10-856-499-1049	Sequence 1049, Ap
11	49	51.6	368	6 US-11-097-143-36393	Sequence 16393, A
12	49	51.6	382	5 US-10-732-923-6514	Sequence 6514, Ap
13	49	51.6	332	4 US-10-424-599-214342	Sequence 214342,
14	49	51.6	392	5 US-10-739-930-9372	Sequence 9372, Ap
15	49	51.6	399	5 US-10-732-923-6515	Sequence 6515, Ap
16	49	51.6	399	5 US-10-732-923-6517	Sequence 6517, Ap
17	49	51.6	399	5 US-10-732-923-6518	Sequence 6518, Ap
18	49	51.6	399	5 US-10-732-923-6519	Sequence 6519, Ap
19	49	51.6	408	5 US-10-732-923-6586	Sequence 6586, Ap
20	49	51.6	410	5 US-10-732-923-6592	Sequence 6592, Ap
21	49	51.6	418	4 US-10-425-114-39285	Sequence 39285, A
22	49	51.6	459	4 US-10-168-273B-2	Sequence 2, Appl1
23	49	51.6	459	4 US-10-437-863-13793	Sequence 13793,
24	49	51.6	459	5 US-10-732-923-6572	Sequence 6572, Ap
25	49	51.6	459	5 US-10-732-923-6599	Sequence 6599, Ap
26	49	51.6	459	5 US-10-732-923-6600	Sequence 6600, Ap
27	49	51.6	686	5 US-10-732-923-6592	Sequence 6592, Ap

28	48	50.5	157	4 US-10-437-963-112319	Sequence 112319,
29	48	50.5	157	4 US-10-425-115-276663	Sequence 276663,
30	48	50.5	176	4 US-10-767-701-57790	Sequence 57790, A
31	48	50.5	213	5 US-10-732-923-6653	Sequence 6653, Ap
32	48	50.5	294	4 US-10-425-115-187125	Sequence 187125,
33	48	50.5	372	5 US-10-732-923-6655	Sequence 6655, Ap
34	48	50.5	401	3 US-09-533-029-80	Sequence 80, Appl1
35	48	50.5	401	5 US-10-732-923-6478	Sequence 6478, Ap
36	48	50.5	401	5 US-10-732-923-6481	Sequence 6481, Ap
37	48	50.5	402	5 US-10-732-923-6467	Sequence 6467, Ap
38	48	50.5	427	4 US-10-412-998B-192	Sequence 192, App
39	48	50.5	466	3 US-09-934-455-72	Sequence 72, Appl1
40	48	50.5	466	4 US-10-225-066A-1026	Sequence 1026, Ap
41	48	50.5	466	4 US-10-225-067-110	Sequence 110, App
42	48	50.5	466	4 US-10-374-780A-2280	Sequence 2280, Ap
43	48	50.5	466	5 US-10-732-923-6491	Sequence 6491, Ap
44	48	50.5	466	5 US-10-225-066A-1026	Sequence 1026, Ap
45	48	50.5	468	5 US-10-732-923-6547	Sequence 6547, Ap
46	48	50.5	470	4 US-10-425-114-41795	Sequence 41795, A
47	48	50.5	477	4 US-10-424-599-228575	Sequence 228575,
48	48	50.5	497	4 US-10-437-963-185620	Sequence 185620,
49	48	50.5	519	4 US-10-094-749-1972	Sequence 1972, Ap
50	48	50.5	670	4 US-10-108-260A-3103	Sequence 3103, Ap
51	48	50.5	684	6 US-11-097-143-14865	Sequence 14865, A
52	48	50.5	834	5 US-10-732-923-6492	Sequence 6492, Ap
53	48	50.5	93	4 US-10-425-115-226040	Sequence 226040,
54	47	49.5	124	5 US-10-450-763-43238	Sequence 43238, A
55	47	49.5	265	4 US-10-425-114-62293	Sequence 62293, A
56	47	49.5	270	4 US-10-425-114-70345	Sequence 70345, A
57	47	49.5	470	6 US-11-097-143-9846	Sequence 9846, Ap
58	47	49.5	517	6 US-11-097-143-10527	Sequence 10527, A
59	47	49.5	528	4 US-10-437-963-175792	Sequence 175792,
60	47	49.5	639	5 US-10-741-849-7206	Sequence 7206, Ap
61	47	49.5	645	6 US-11-097-143-11427	Sequence 11427, A
62	47	49.5	697	6 US-11-097-143-31923	Sequence 31923, A
63	47	49.5	58	3 US-09-864-761-48985	Sequence 48985, A
64	46	48.4	109	5 US-10-637-313-8	Sequence 8, Appl1
65	46	48.4	109	5 US-10-637-313-48	Sequence 48, Appl1
66	46	48.4	125	5 US-10-507-734-26	Sequence 26, Appl1
67	46	48.4	126	4 US-10-425-114-41644	Sequence 41644, A
68	46	48.4	127	4 US-10-767-701-40532	Sequence 40532, A
69	46	48.4	206	4 US-10-425-115-252726	Sequence 252726,
70	46	48.4	233	5 US-10-637-313-12	Sequence 12, Appl1
71	46	48.4	243	5 US-10-637-313-50	Sequence 50, Appl1
72	46	48.4	277	6 US-11-097-143-37026	Sequence 37026, A
73	46	48.4	292	4 US-10-282-122A-70106	Sequence 70106, A
74	46	48.4	305	5 US-10-450-763-51459	Sequence 51459, A
75	46	48.4	315	4 US-10-424-599-220979	Sequence 220979,
76	46	48.4	325	5 US-10-389-771-7	Sequence 7, Appl1
77	46	48.4	363	4 US-10-282-122A-49514	Sequence 49514, A
78	46	48.4	415	5 US-10-637-313-26	Sequence 26, Appl1
79	46	48.4	415	5 US-10-893-315-101	Sequence 101, App
80	46	48.4	579	5 US-10-893-315-105	Sequence 105, App
81	46	48.4	589	6 US-11-097-143-3579	Sequence 3579, Ap
82	46	48.4	615	4 US-10-162-335-72	Sequence 72, Appl1
83	46	48.4	615	5 US-10-637-313-22	Sequence 22, Appl1
84	46	48.4	616	5 US-10-637-313-14	Sequence 14, Appl1
85	46	48.4	621	5 US-10-637-313-16	Sequence 16, Appl1
86	46	48.4	621	5 US-10-637-313-18	Sequence 18, Appl1
87	46	48.4	621	5 US-10-637-313-14	Sequence 14, Appl1
88	46	48.4	622	5 US-10-637-313-18	Sequence 18, Appl1
89	46	48.4	626	5 US-10-307-734-25	Sequence 25, Appl1
90	46	48.4	630	5 US-10-732-923-2689	Sequence 2689, Ap
91	46	48.4	644	4 US-10-162-335-74	Sequence 74, Appl1
92	46	48.4	644	4 US-10-162-335-84	Sequence 84, Appl1
93	46	48.4	644	5 US-10-637-313-2	Sequence 2, Appl1
94	46	48.4	644	5 US-10-637-313-4	Sequence 4, Appl1
95	46	48.4	644	5 US-10-637-313-16	Sequence 16, Appl1
96	46	48.4	644	5 US-10-637-313-52	Sequence 52, Appl1
97	46	48.4	644	5 US-10-637-313-54	Sequence 54, Appl1
98	46	48.4	644	5 US-10-637-313-56	Sequence 56, Appl1
99	46	48.4	644	5 US-10-637-313-58	Sequence 58, Appl1
100	46	48.4	644	5 US-10-637-313-58	Sequence 58, Appl1

## ALIGNMENTS

RESULT 1  
US-09-858-332-2  
Sequence 2, Application US/09858332  
Patent No. US20020164718A1  
GENERAL INFORMATION:  
APPLICANT: Tchaga, Grigory S.  
APPLICANT: Jokhadze, George  
TITLE OF INVENTION: Metal Ion Affinity Tags and Methods for  
TITLE OF INVENTION: Using the Same  
FILE REFERENCE: CLON056CIP  
CURRENT APPLICATION NUMBER: US/09/858,332  
CURRENT FILING DATE: 2002-07-02  
PRIOR APPLICATION NUMBER: 09/404,017  
PRIOR FILING DATE: 1999-09-23  
PRIOR APPLICATION NUMBER: 60/101,867  
PRIOR FILING DATE: 1998-09-25  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO. 2  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic peptide  
US-09-858-332-2

Query Match 100.0%; Score 95; DB 3; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1,6e-06;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLHNVKKEHHAAHN 16  
|||||

DB 1 HLHNVKKEHHAAHN 16

RESULT 2  
US-10-762-588-1  
Sequence 1, Application US/10762588  
Publication No. US20040180415A1  
GENERAL INFORMATION:  
APPLICANT: Tchaga, Grigory S.  
APPLICANT: Jokhadze, George G.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR PROTEIN  
TITLE OF INVENTION: PURIFICATION  
FILE REFERENCE: CLON-056US2  
CURRENT APPLICATION NUMBER: US/10/762,588  
CURRENT FILING DATE: 2004-01-21  
PRIOR APPLICATION NUMBER: 09/858,332  
PRIOR FILING DATE: 2001-05-15  
PRIOR APPLICATION NUMBER: 60/441,804  
PRIOR FILING DATE: 2003-01-21  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Synthetic Affinity Peptide  
US-10-762-588-1

Query Match 100.0%; Score 95; DB 4; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1,6e-06;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLHNVKKEHHAAHN 16  
|||||

DB 1 HLHNVKKEHHAAHN 16

RESULT 3  
US-10-748-055-1  
Sequence 1, Application US/10748055  
Publication No. US20040214292A1  
GENERAL INFORMATION:  
APPLICANT: MOTODA, Yoko et al  
TITLE OF INVENTION: METHOD OF PRODUCING TEMPLATE DNA AND METHOD OF PRODUCING PROTEIN  
TITLE OF INVENTION: FREE PROTEIN SYNTHESIS SYSTEM USING THE SAME  
FILE REFERENCE: 1686-0108P  
CURRENT APPLICATION NUMBER: US/10/748,055  
CURRENT FILING DATE: 2003-12-31  
PRIOR APPLICATION NUMBER: PCT/JP02/06261  
PRIOR FILING DATE: 2002-06-24  
PRIOR APPLICATION NUMBER: JP P2001-201356  
PRIOR FILING DATE: 2001-07-02  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 20  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: native His tag  
US-10-748-055-1

Query Match 100.0%; Score 95; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2e-06;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLHNVKKEHHAAHN 16  
|||||

DB 4 HLHNVKKEHHAAHN 19

RESULT 4  
US-09-858-332-15  
Sequence 15, Application US/09858332  
Patent No. US20020164718A1  
GENERAL INFORMATION:  
APPLICANT: Tchaga, Grigory S.  
APPLICANT: Jokhadze, George  
TITLE OF INVENTION: Metal Ion Affinity Tags and Methods for  
TITLE OF INVENTION: Using the Same  
FILE REFERENCE: CLON056CIP  
CURRENT APPLICATION NUMBER: US/09/858,332  
CURRENT FILING DATE: 2002-07-02  
PRIOR APPLICATION NUMBER: 09/404,017  
PRIOR FILING DATE: 1999-09-23  
PRIOR APPLICATION NUMBER: 60/101,867  
PRIOR FILING DATE: 1998-09-25  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 279  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic peptide  
US-09-858-332-15

Query Match 100.0%; Score 95; DB 3; Length 279;  
Best Local Similarity 100.0%; Pred. No. 2,8e-05;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLHNVKKEHHAAHN 16  
|||||

DB 11 HLHNVKKEHHAAHN 26

RESULT 5  
US-10-634-645-14

```
; Sequence 14, Application US/10634645
; Publication No. US20050032173A1
; GENERAL INFORMATION:
; APPLICANT: Rojas, Mauricio
; APPLICANT: Mora, Ana L.
; TITLE OF INVENTION: FUSION PROTEINS WITH A MEMBRANE TRANSLOCATING SEQUENCE AND METHOD
; TITLE OF INVENTION: USING SAME TO INHIBIT AN IMMUNE RESPONSE
; FILE REFERENCE: 60068.2US01
; CURRENT APPLICATION NUMBER: US/10/634,645
; CURRENT FILING DATE: 2003-08-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Tag sequence
US-10-634-645-14

Query Match          91.6%; Score 87; DB 5; Length 19;
Best Local Similarity 93.8%; Pred. No. 2,4e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLHNVKKEFHAAH 16
    |||||
Db 3 HLHNVKKEFHAAH 18

RESULT 6
US-11-097-143-8130
; Sequence 8130, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: C1000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 8130
; LENGTH: 837
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-8130

Query Match          55.8%; Score 53; DB 6; Length 837;
Best Local Similarity 53.3%; Pred. No. 52;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 HLHNVKKEFHAAH 15
    |||||
Db 318 HHVHAHAHAHAH 332
```

```
RESULT 7
US-11-097-143-20379
; Sequence 20379, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: C1000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 20379
; LENGTH: 402
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-20379

Query Match          53.7%; Score 51; DB 6; Length 402;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 HNVKKEFHAAH 15
    |||||
Db 239 HHVHAHAHAHAH 250

RESULT 8
US-10-425-115-344916
; Sequence 344916, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 344916
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(565)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_77730C.1.pep
```

US-10-425-115-344916

Query Match	53.7%;	Score 51;	DB 4;	Length 565;
Best Local Similarity	53.3%;	Pred. No. 67;		
Matches	8;	Conservative	4;	Mismatches 3;
				Indels

OY	2	L I H N V H K E E H A L A H N	16
		: :   :   :	:
D b	113	L S H G V H Q E K H S H A Q S	127

**RESULT 9**  
**ITS-10-42**

```

US-10-425-115-365836,1
; Sequence 365836, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 365836
; LENGTH: 60
; TYPR: PRT
; ORGANISM: Zea mays
; FEATURES:
; OTHER INFORMATION: Clone ID: MRT4577_96807C.1.pep
US-10-425-115-365836

```

Query Match	52.6%	Score 50;	DB 4;	Length 60;
Best Local Similarity	50.0%;	Pred. No. 9.6;		
Matches	8;	Conservative	2;	Mismatches 6;
				Indels 0;
				Gaps 0

Qy	1	HLIHNVHKEEHAHAHN	16
		: :	
Db	30	HLVDNIHCTEHLAKGN	45

RESULT 10  
US-10-856-499-1049

```

1  Sequence 1049, Application US/10856499
2  Publication NO. US20040259145A1
3  GENERAL INFORMATION:
4  APPLICANT: Wood, Marion
5  APPLICANT: Shenk, Michael A.
6  APPLICANT: McGrath, Annette
7  APPLICANT: Glenn, Matthew
8  TITLE OF INVENTION: Compositions and Methods for the
9  TITLE OF INVENTION: Modification of Gene Transcription
10 FILE REFERENCE: 11000.102162
11 CURRENT APPLICATION NUMBER: US/10/856,499
12 CURRENT FILING DATE: 2004-05-28
13 NUMBER OF SEQ ID NOS: 2370
14 SOFTWARE: FastSeq for Windows Version 4.0
15 SEQ ID NO 1049
16     LENGTH: 134
17     TYPE: PR1
18 ORGANISM: Pinus radiata
19 US-10-856-499-1049

```

Query Match	51.6%	Score 49;	DB 5;	Length 134;
Best Local Similarity	37.5%	Pred. No. 30;		
Matches	6;	Conservative	6;	Mismatches 4;
				Indels 0;
				Gaps 0

```
QY      1 HLHNVHKEEHAHAHN 16
          ||:|:|:|:|:|:|:
Db      96 HLKNITHRRKPVSHS 111
```

RESULT 11

US-11-097-143-36393  
 ; Sequence 36393, Application US/11097143  
 ; Publication NO. US2005020855A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Venter, J. Craig  
 ; APPLICANT: et al.  
 ; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
 ; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
 ; TITLE OF INVENTION: DROSOPHILA GENES.

Query Match	51.6%	Score 49	DB 6	Length 368
Best Local Similarity	42.9%	Pred. No. 82		
Matches 6; Conservative	4	Mismatches	0	Gaps 0

QY	3	IHNVKKEEHAHNN	16
		: : : :	
Db	258	VHSHTHSHSHSHN	27

```

RESULT 12
US-10-732-923-6514
; Sequence 6514, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15152796/C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 6514
; LENGTH: 382
; TYPE: PRT
; ORGANSIM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(382)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-732-923-6514

```

Query Match	51.6%	Score 49;	DB 5;	Length 382;
Best Local Similarity	42.9%;	Pred. NO. 85;		
Matches 6; Conservative	5;	Mismatches 3;	Indels 0;	Gaps 0



Qy 1 HLHNHVKKEEHAH 14  
 ||:|:|:|:|:  
 Db 96 HLKNIHRRKPVSHS 111

## RESULT 13

US-10-424-599-214342  
 ; Sequence 214342, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 214342  
 ; LENGTH: 392  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT1847\_35579C.1.pep  
 US-10-424-599-214342

Query Match 51.6%; Score 49; DB 4; Length 392;  
 Best Local Similarity 37.5%; Pred. No. 87;  
 Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 HLHNHVKKEEHAH 16  
 ||:|:|:|:|:  
 Db 96 HLKNIHRRKPVSHS 111

## RESULT 14

US-10-739-930-9372  
 ; Sequence 9372, Application US/10739930  
 ; Publication No. US20040216190A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kovalic, David K.  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
 ; FILE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT  
 ; FILE REFERENCE: 38-21(53377)B  
 ; CURRENT APPLICATION NUMBER: US/10/739,930  
 ; CURRENT FILING DATE: 2003-12-18  
 ; NUMBER OF SEQ ID NOS: 11088  
 ; SEQ ID NO 9372  
 ; LENGTH: 392  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(392)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: GLVMA-23APR03-C790\_2.p  
 US-10-739-930-9372

Query Match 51.6%; Score 49; DB 5; Length 392;  
 Best Local Similarity 37.5%; Pred. No. 87;  
 Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 HLHNHVKKEEHAH 16  
 ||:|:|:|:|:  
 Db 96 HLKNIHRRKPVSHS 111

RESULT 15  
 US-10-732-923-6515

; Sequence 6515, Application US/10732923  
 ; Publication No. US20050108791A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edgeton, Michael D  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
 ; FILE REFERENCE: 38-15(52796)C  
 ; CURRENT APPLICATION NUMBER: US/10/732,923  
 ; CURRENT FILING DATE: 2003-12-10  
 ; PRIOR APPLICATION NUMBER: 10/310,154  
 ; PRIOR FILING DATE: 2002-12-04  
 ; NUMBER OF SEQ ID NOS: 24149  
 ; SEQ ID NO 6515  
 ; LENGTH: 399  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(399)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 US-10-732-923-6515

Query Match 51.6%; Score 49; DB 5; Length 399;  
 Best Local Similarity 37.5%; Pred. No. 89;  
 Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 HLHNHVKKEEHAH 16  
 ||:|:|:|:|:  
 Db 96 HLKNIHRRKPVSHS 111

## RESULT 16

US-10-732-923-6517  
 ; Sequence 6517, Application US/10732923  
 ; Publication No. US20050108791A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edgeton, Michael D  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
 ; FILE REFERENCE: 38-15(52796)C  
 ; CURRENT APPLICATION NUMBER: US/10/732,923  
 ; CURRENT FILING DATE: 2003-12-10  
 ; PRIOR APPLICATION NUMBER: 10/310,154  
 ; PRIOR FILING DATE: 2002-12-04  
 ; NUMBER OF SEQ ID NOS: 24149  
 ; SEQ ID NO 6517  
 ; LENGTH: 399  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(399)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 US-10-732-923-6517

Query Match 51.6%; Score 49; DB 5; Length 399;  
 Best Local Similarity 37.5%; Pred. No. 89;  
 Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 HLHNHVKKEEHAH 16  
 ||:|:~|:~|:~|:  
 Db 96 HLKNIHRRKPVSHS 111

## RESULT 17

US-10-732-923-6518  
 ; Sequence 6518, Application US/10732923  
 ; Publication No. US20050108791A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edgeton, Michael D  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
 ; FILE REFERENCE: 38-15(52796)C  
 ; CURRENT APPLICATION NUMBER: US/10/732,923  
 ; CURRENT FILING DATE: 2003-12-10  
 ; PRIOR APPLICATION NUMBER: 10/310,154

RESULT 22  
US-10-168-273B-2  
; Sequence 2, Application US/10168273B



Db 102 HLKNIHRRKPVHSHS 117

```
RESULT 27
US-10-732-923-6592
; Sequence 6592, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 6592
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-732-923-6592
```

Query Match 51.6%; Score 49; DB 5; Length 686;  
Best Local Similarity 37.5%; Pred. No. 1.5e+02;  
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 HLHNHVKESHAAH 16  
||:|:|:|:|:|:|:  
Db 329 HLKNIHRRKPVHSHS 344

```
RESULT 28
US-10-437-963-112139
; Sequence 112139, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 112139
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(157)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_16052C.1.pep
US-10-437-963-112139
```

Query Match 50.5%; Score 48; DB 4; Length 157;  
Best Local Similarity 46.7%; Pred. No. 48;  
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 HLHNHVKESHAAH 15  
||:|:|:|:|:|:|:  
Db 101 HYKHNAGQOQHHRH 115

RESULT 29  
US-10-425-115-276663

```
; Sequence 276663, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 276663
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(157)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_183900C.1.pep
US-10-425-115-276663
```

Query Match 50.5%; Score 48; DB 4; Length 157;  
Best Local Similarity 70.0%; Pred. No. 48;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 6 VKEESHAAH 15  
:|:|:|:|:|:|:  
Db 50 IHREHAAH 59

```
RESULT 30
US-10-767-701-57790
; Sequence 57790, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 57790
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(176)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: 30972058.pep
US-10-767-701-57790
```

Query Match 50.5%; Score 48; DB 4; Length 176;  
Best Local Similarity 37.5%; Pred. No. 53;  
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 HLHNHVKESHAAH 16  
||:|:|:|:|:|:|:  
Db 109 HLKNIHRRKPVHSHS 124

RESULT 31  
US-10-732-923-6653  
; Sequence 6653, Application US/10732923  
; Publication No. US20050108791A1

```

GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
FILE REFERENCE: 38-15 (527967C
CURRENT APPLICATION NUMBER: US/10/732,923
CURRENT FILING DATE: 2003-12-10
PRIORITY APPLICATION NUMBER: 10/310,154
PRIORITY FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 24149
SEQ ID NO 6653
LENGTH: 213
TYPE: PRT
ORGANISM: Zea mays
FEATURES:
NAME/KEY: unsure
LOCATION: (1)..(213)
OTHER INFORMATION: unsure at all Xaa locations
US-10-732-923-6653

```

Query Match	50.5%	Score 48;	DB 5;	Length 213;
Best Local Similarity	37.5%	Pred. No. 65;		
Matches	6;	Conservative	6;	Mismatches 4;
				Indels 0;
				Gaps 0;

```
QY      1 HLHNVHKEEHAHNN 16
          ||:|:|:|:|:|:
Db     107 HLLKNTHRRKPIHS 122
```

```

RESULT 32
US-10-425-115-187125
Sequence 187125, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 187125
LENGTH: 294
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (1) ..(294)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_102245C.1 pep
US-10-425-115-187125

```

Query Match	50.5%	Score 48;	DB 4;	Length 294;
Best Local Similarity	37.5%;	Pred. No. 90;		
Matches 6;	Conservative 6;	Mismatches 4;	Indels 0;	Gaps 0;

```
OY      1 HLHNVHKEEHAHAN 16  
        ||:||:|:|:  
Db     107 HLLKNIHRRKPIHS 122
```

RESULT 33  
US-10-732-923-6655  
; Sequence 6655, Application US/107332923  
; Publication NO. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923

```

: : CURRENT FILING DATE: 2003-12-10
: : PRIOR APPLICATION NUMBER: 10/310,154
: : PRIOR FILING DATE: 2002-12-04
: : NUMBER OF SEQ ID NOS: 24149
: : SEQ ID NO 6655
: : LENGTH: 372
: : TYPE: PRT
: : ORGANISM: Zea mays
: : FEATURE:
: : NAME/KEY: unsure
: : LOCATION: (1) ..(372)
: : OTHER INFORMATION: unsure at all Xaa locations
US-10-732-923-6655

Query Match          50.5%; Score 48; DB 5; Length 372;
Best Local Similarity 37.5%; Pred. No. 1,1e+02;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```

Qy	1	HLIHNVAKEEHAAAHN	16
		: : : : :	
Db	138	HLKNIHRRKPISHS	153

RESULT 34  
US-09-533-029-80

```

Publication No.: US20030046722A1
GENERAL INFORMATION:
APPLICANT: Heard, Jacqueline
APPLICANT: Brown, Pierre
APPLICANT: Riechmann, Josef-Luis
APPLICANT: Keddie, James
APPLICANT: Pineda, Omaidra
APPLICANT: Adam, Luc
APPLICANT: Samaha, Raymond
APPLICANT: Zhang, James
APPLICANT: Yu, Guo-Liang
APPLICANT: Ratcliffe, Oliver
APPLICANT: Pilgrim, Marsha
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Reuber, Lynne
TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDE
FILE REFERENCE: MEI-010
CURRENT APPLICATION NUMBER: US/09/533,029
CURRENT FILING DATE: 2000-03-22
EARLIER APPLICATION NUMBER: 60/125,814
EARLIER FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 121
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 80
LENGTH: 401
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G261
US-09-533-029-80

```

Query Match	50.5%	Score 48;	DB 3;	Length 401;
Best local Similarity	37.5%	Pred. No. 1.2e+02;		
Matches 6;	Conservative 6;	Mismatches 4;	Indels 0;	Gaps 0;

QY 1 HLIHNVHKEEHAHAHN 16  
||:|:|:|:|:|:  
Db 97 HLMKNIHRRKPPVSHS 112

RESULT 35  
US-10-732-923-6478  
; Sequence 6478, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES

```

FILE REFERENCE: 38-15(52796) C
CURRENT APPLICATION NUMBER: 10/10/732,923
CURRENT FILING DATE: 2003-12-10
PRIOR APPLICATION NUMBER: 10/310,154
PRIOR FILING DATE: 2002-13-04
NUMBER OF SEQ ID NOS: 24149
SEQ ID NO 6478
LENGTH: 401
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-10-732-923-6478

```

Query Match	50.5%	Score 48	DB 5	Length 401
Best Local	37.5%	Pred. No. 1.2e+02		
Matches	6	Conservative	4	Indels 0
				Gaps 0

Qy 1 HLHNVHKEEHAHAHN 16  
||:|:|:|:|:|:|:  
Db 97 HLMKNIHRRKPVTHSHS 112

```

RESULT 36
US-10-732-923-6481
Sequence 6481, Application US/10732923
Publication No. US20050108791A1
GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPE
FILE REFERENCE: 38-15 (52796)C
CURRENT APPLICATION NUMBER: US/10732,923
CURRENT FILING DATE: 2003-12-10
PRIOR APPLICATION NUMBER: 10/310,154
PRIOR FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 24149
SEQ ID NO 6481
LENGTH: 401
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-10-732-923-6481

```

Query Match	50.5%	Score	48	DB	5	Length	401
Best Local Similarity	37.5%	Pred.	No.	1.2e+02			
Matches	6	Conservative	6	Mismatches	4	Indels	0
				Gaps	0		0

QY	1	HLIHNVHKEEHAHAHN	16
		: : : : : :	
Db	97	HLMKNIHRRKPVHSHS	112

RESULT 37  
US-10-732-923-6467  
; Sequence 6467, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPEPRESS  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10732, 923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 6467  
; LENGTH: 402  
; TYPE: PRT  
; ORGANISM: Phaseolus acutifolius  
; US-10-732-923-6467

Query Match	50.5%	Score 48	DB 5	Length 402
Best Local Similarity	37.5%	Pred. No. 1.2e+02		
Matches	6	Conservative	4	Indels 0
				Gaps 0

QY 1 HLJHNVHKEEHAHVN 16

Db 95 HLMKNIHRRKPPVSHS 110

```

RESULT 38
US-10-732-923-6567
; Sequence 6567, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPIC
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 6567
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Medicago sativa
; US-10-732-923-6567

```

Query Match 50.5%; Score 48; DB 5; Length 402;  
Best Local Similarity 37.5%; Pred. NO. 1.2e+02;  
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY	1	HLIHNVHKEEHAHAHN	16
		: : : : : :	
Db	94	HLMKNIHRRKPVHSHS	109

RESULT 39  
US-10-412-6998-192  
Sequence 192, Application US/10412699B  
Publication No. US20040045049A1  
GENERAL INFORMATION:  
APPLICANT: Mendel Biotechnology, Inc.  
APPLICANT: Zhang, James  
APPLICANT: Fromm, Michael E.  
APPLICANT: Heard, Jacqueline E.  
APPLICANT: Riechmann, Jose Luis  
APPLICANT: Adam, Luc J.  
APPLICANT: Biron, Pierre E.  
APPLICANT: Pineda, Omalra  
APPLICANT: Reuber, T. Lynne  
APPLICANT: Keddie, James S.  
APPLICANT: Yu, Guo-Liang  
APPLICANT: Jiang, Cai-Zhong  
APPLICANT: Samaha, Raymond R.  
APPLICANT: Pilgrim, Marsha L.  
APPLICANT: Creelman, Robert A.  
APPLICANT: Dubell, Arnold N.  
APPLICANT: Ratcliffe, Oliver  
APPLICANT: Kumimoto, Roderick  
APPLICANT: Sherman, Bradley K.  
TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants  
FILE REFERENCE: MBI-004861P  
CURRENT APPLICATION NUMBER: US/10/412,699B  
CURRENT FILING DATE: 2003-04-10  
PRIOR APPLICATION NUMBER: 09/394,519  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: 09/489,376  
PRIOR FILING DATE: 2000-01-21  
PRIOR APPLICATION NUMBER: 09/506,720  
PRIOR FILING DATE: 2000-02-17  
PRIOR APPLICATION NUMBER: 09/533,030  
PRIOR FILING DATE: 2000-03-22  
PRIOR APPLICATION NUMBER: 09/533,392  
PRIOR FILING DATE: 2000-03-22  
PRIOR APPLICATION NUMBER: 09/533,029  
PRIOR FILING DATE: 2000-03-22  
PRIOR APPLICATION NUMBER: 09/532,591

```

; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 192
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G261
US-10-412-6998-192

```

```

Query Match      50.5%; Score 48; DB 4; Length 427;
Best Local Similarity 37.5%; Pred. No. 1.3e+02;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy      1 HLHNHVKKEHAAHNN 16
Db      123 HLMKNTHRRKPVSHS 138

```

```

RESULT 40
US-09-934-455-72
; Sequence 72, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddie, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pineda, Omaira
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: MBI-0022
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: MBI-0023
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-934-455-72

```

```

Query Match      50.5%; Score 48; DB 3; Length 466;
Best Local Similarity 37.5%; Pred. No. 1.4e+02;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy      1 HLHNHVKKEHAAHNN 16
Db      105 HLMKNTHRRKPVSHS 120

```

(oldsn) 2/10/2011 11:11



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2006, 13:15:13 ; Search time 50 Seconds  
(without alignments)  
26,456 Million cell updates/sec

Title: US-10-762-588-1

Perfect score: 95  
Sequence: 1 HLIHNVKKEHAAAHN 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 100 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/6\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/H\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/PCUS\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/PCUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	51.6	134	2	US-09-640-211A-1049
2	48	50.5	161	2	US-09-489-039A-9273
3	48	50.5	400	2	US-09-543-681A-6151
4	48	50.5	401	2	US-09-533-029-80
5	47	49.5	491	2	US-09-248-796A-18483
6	46	48.4	47	2	US-09-612-126-4
7	46	48.4	62	2	US-09-612-126-7
8	46	48.4	83	2	US-09-612-126-6
9	46	48.4	94	2	US-09-612-126-10
10	46	48.4	179	2	US-09-612-126-11
11	46	48.4	186	2	US-09-612-126-8
12	46	48.4	255	2	US-09-612-126-1
13	46	48.4	579	2	US-09-949-002-475
14	46	48.4	579	2	US-09-949-002-481
15	45	47.4	25	2	US-09-721-154-14
16	45	47.4	353	2	US-09-328-352-4930
17	45	47.4	367	2	US-09-540-236-2996
18	45	47.4	501	2	US-09-489-039A-9286
19	44.5	46.8	289	2	US-09-540-236-2019
20	44	46.3	115	2	US-09-991-181-95
21	44	46.3	115	2	US-09-990-444-95
22	44	46.3	115	2	US-09-997-333-95
23	44	46.3	115	2	US-09-992-598-95
24	44	46.3	175	2	US-10-104-047-3197
25	44	46.3	175	2	US-09-640-211A-764
26	44	46.3	316	2	US-09-801-876B-7
27	44	46.3	316	2	US-09-252-991A-27084
28	44	46.3	316	2	US-10-254-869-7
29	44	46.3	316	2	US-10-667-442-7
30	44	46.3	403	2	US-09-801-876B-5
31	44	46.3	403	2	US-10-254-869-5
32	44	46.3	403	2	US-10-667-442-5
33	44	46.3	404	2	US-09-801-876B-4
34	44	46.3	404	2	US-10-254-869-4
35	44	46.3	404	2	US-10-667-442-4
36	44	46.3	473	2	US-09-248-796A-16084
37	44	46.3	480	1	US-07-882-292-2
38	44	46.3	480	1	US-08-331-644-2
39	44	46.3	480	4	PCT-US93-04102-2
40	44	46.3	570	2	US-10-104-047-2718
41	43.5	45.8	224	2	US-09-302-540-12716
42	43.5	45.8	834	1	US-08-861-464-4
43	43.5	45.8	834	1	US-08-996-001-4
44	43.5	45.8	834	2	US-09-323-433A-4
45	43.5	45.8	834	2	US-09-826-752-4
46	43	45.3	29	2	US-09-354-231B-51
47	43	45.3	29	2	US-09-128-602B-51
48	43	45.3	29	2	US-09-995-297-2
49	43	45.3	106	2	US-09-248-796A-25254
50	43	45.3	198	2	US-09-270-767-43398
51	43	45.3	213	2	US-09-248-796A-16185
52	43	45.3	217	2	US-09-248-796A-14838
53	43	45.3	218	2	US-09-252-991A-25291
54	43	45.3	322	2	US-08-956-970A-2
55	43	45.3	355	1	US-08-458-182-17
56	43	45.3	357	2	US-09-723-806A-9
57	43	45.3	398	2	US-09-461-474-17
58	43	45.3	453	1	US-08-244-205-13
59	43	45.3	453	2	US-09-161-994A-10
60	43	45.3	453	4	PCT-US92-10284-13
61	43	45.3	477	2	US-09-302-540-11909
62	43	45.3	615	2	US-09-543-681A-8049
63	43	45.3	850	2	US-09-893-525-42
64	42.5	44.7	491	2	US-10-029-180-106
65	42	44.2	65	2	US-09-248-796A-21174
66	42	44.2	115	2	US-09-461-325-160
67	42	44.2	115	2	US-10-012-542-160
68	42	44.2	321	2	US-10-115-123-160
69	42	44.2	321	2	US-09-543-681A-4740
70	42	44.2	527	2	US-09-270-767-46469
71	42	44.2	587	2	US-09-252-991A-22704
72	42	44.2	686	2	US-09-949-016-11203
73	42	44.2	696	2	US-09-351-414-2
74	42	44.2	832	2	US-09-634-252A-4
75	42	44.2	1284	2	US-10-396-144-5
76	42	44.2	2353	2	US-08-384-709A-50
77	41.5	43.7	342	2	US-09-134-001C-4190
78	41.5	43.7	431	1	US-08-911-023-2
79	41.5	43.7	755	2	US-09-642-034-5
80	41	43.2	61	2	US-09-248-796A-25020
81	41	43.2	84	2	US-09-673-395A-182
82	41	43.2	92	2	US-09-543-681A-6724
83	41	43.2	114	2	US-09-248-796A-22116
84	41	43.2	162	1	US-08-624-125-8
85	41	43.2	162	1	US-08-624-125-8
86	41	43.2	162	1	US-08-663-310-4
87	41	43.2	162	1	US-08-663-310-11
88	41	43.2	162	1	US-09-006-491-1
89	41	43.2	162	1	US-09-006-491-11
90	41	43.2	162	2	US-09-335-919-4
91	41	43.2	162	2	US-09-335-919-11
92	41	43.2	162	2	US-08-980-832-34
93	41	43.2	162	2	US-08-937-155-5
94	41	43.2	162	2	US-08-937-155-8
95	41	43.2	162	2	US-09-323-998B-5
96	41	43.2	162	2	US-09-323-998B-8
97	41	43.2	162	2	US-09-920-923B-34
98	41	43.2	171	2	US-09-640-211A-1065
99	41	43.2	187	2	US-09-396-937-8
100	41	43.2	203	2	US-09-270-767-35326

```

RESULT 3
US-09-543-681A-6151
/ Sequence 6151, Application US/09543681A
/ Patent No. 6605709
/ GENERAL INFORMATION:
/ APPLICANT: GARY BRETON
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
/ TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

```

```

RESULT 5
US-09-248-796A-18483
; Sequence 18483, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132

```

CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 18483  
LENGTH: 491  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-18483

Query Match 49.5%; Score 47; DB 2; Length 491;  
Best Local Similarity 46.7%; Pred. No. 26;  
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
QY 1 HLHNHVKKEHAAH 15  
DB 385 HVLNHHKHHHHHH 399

RESULT 6  
US-09-612-126-4  
Sequence 4, Application US/09612126  
Patent No. 6284726  
GENERAL INFORMATION:  
APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF  
TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH  
FILE REFERENCE: 6056-258 CT1  
CURRENT APPLICATION NUMBER: US/09/612,126  
CURRENT FILING DATE: 2000-07-07  
PRIOR APPLICATION NUMBER: 60/107,844  
PRIOR FILING DATE: 1998-11-10  
PRIOR APPLICATION NUMBER: PCT/US99/26377  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 47  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Human high  
OTHER INFORMATION: molecular weight kininogen light chain amino acids  
OTHER INFORMATION: Gly(456) through Lys(502)  
US-09-612-126-4

Query Match 48.4%; Score 46; DB 2; Length 47;  
Best Local Similarity 40.0%; Pred. No. 3.1;  
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
QY 1 HLHNHVKKEHAAH 15  
DB 16 HVLNHHKHHHHHH 30

RESULT 7  
US-09-612-126-7  
Sequence 7, Application US/09612126  
Patent No. 6284726  
GENERAL INFORMATION:  
APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF  
TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH  
FILE REFERENCE: 6056-258 CT1  
CURRENT APPLICATION NUMBER: US/09/612,126  
CURRENT FILING DATE: 2000-07-07  
PRIOR APPLICATION NUMBER: 60/107,844  
PRIOR FILING DATE: 1998-11-10  
PRIOR APPLICATION NUMBER: PCT/US99/26377  
PRIOR FILING DATE: 1999-11-09

NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 7  
LENGTH: 62  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Human high  
OTHER INFORMATION: molecular weight kininogen light chain amino acids  
OTHER INFORMATION: His(441) through Lys(502)  
US-09-612-126-7

Query Match 48.4%; Score 46; DB 2; Length 62;  
Best Local Similarity 40.0%; Pred. No. 4.1;  
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
QY 1 HLHNHVKKEHAAH 15  
DB 31 HVLNHHKHHHHHH 45

RESULT 8  
US-09-612-126-6  
Sequence 6, Application US/09612126  
Patent No. 6284726  
GENERAL INFORMATION:  
APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF  
TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH  
FILE REFERENCE: 6056-258 CT1  
CURRENT APPLICATION NUMBER: US/09/612,126  
CURRENT FILING DATE: 2000-07-07  
PRIOR APPLICATION NUMBER: 60/107,844  
PRIOR FILING DATE: 1998-11-10  
PRIOR APPLICATION NUMBER: PCT/US99/26377  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 83  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Human high  
OTHER INFORMATION: molecular weight kininogen light chain amino acids  
OTHER INFORMATION: Lys(420) through Lys(502)  
US-09-612-126-6

Query Match 48.4%; Score 46; DB 2; Length 83;  
Best Local Similarity 40.0%; Pred. No. 5.6;  
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
QY 1 HLHNHVKKEHAAH 15  
DB 52 HVLNHHKHHHHHH 66

RESULT 9  
US-09-612-126-10  
Sequence 10, Application US/09612126  
Patent No. 6284726  
GENERAL INFORMATION:  
APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF  
TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH  
FILE REFERENCE: 6056-258 CT1  
CURRENT APPLICATION NUMBER: US/09/612,126  
CURRENT FILING DATE: 2000-07-07  
PRIOR APPLICATION NUMBER: 60/107,844  
PRIOR FILING DATE: 1998-11-10  
PRIOR APPLICATION NUMBER: PCT/US99/26377  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 12

```
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 10
/ LENGTH: 94
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Human high
/ OTHER INFORMATION: molecular weight kinogen light chain amino acids
/ OTHER INFORMATION: Lys(420) through Ser(513)
US-09-612-126-10
```

```
Query Match          48.4%; Score 46; DB 2; Length 94;
Best Local Similarity 40.0%; Pred. No. 6.4;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 HLHNHVKKEHNAH 15
|:::|::|::|
Db      52 HVLDDGHKHKHGHGH 66
```

```
RESULT 10
US-09-612-126-11
/ Sequence 11, Application US/09612126
/ Patent No. 6284726
```

```
/ GENERAL INFORMATION:
/ APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
/ TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
/ TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAIN 5
/ FILE REFERENCE: 6056-258 CT1
/ CURRENT APPLICATION NUMBER: US/09/612,126
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/107,844
/ PRIOR FILING DATE: 1998-11-10
/ PRIOR FILING DATE: 1999-11-09
/ PRIOR FILING DATE: 1999-11-09
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 11
/ LENGTH: 179
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Human high
/ OTHER INFORMATION: molecular weight kinogen light chain amino acids
/ OTHER INFORMATION: Glu(448) through Ser(626)
US-09-612-126-11
```

```
Query Match          48.4%; Score 46; DB 2; Length 179;
Best Local Similarity 40.0%; Pred. No. 13;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 HLHNHVKKEHNAH 15
|:::|::|::|
Db      24 HVLDDGHKHKHGHGH 38
```

```
RESULT 11
US-09-612-126-8
/ Sequence 8, Application US/09612126
/ Patent No. 6284726
```

```
/ GENERAL INFORMATION:
/ APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
/ TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
/ TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAIN 5
/ FILE REFERENCE: 6056-258 CT1
/ CURRENT APPLICATION NUMBER: US/09/612,126
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/107,844
/ PRIOR FILING DATE: 1998-11-10
/ PRIOR FILING DATE: 1999-11-09
/ PRIOR FILING DATE: 1999-11-09
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn Ver. 2.1
```

```
/ SEQ ID NO 8
/ LENGTH: 186
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Human high
/ OTHER INFORMATION: molecular weight kinogen light chain amino acids
/ OTHER INFORMATION: His(441) through Ser(626)
US-09-612-126-8
```

```
Query Match          48.4%; Score 46; DB 2; Length 186;
Best Local Similarity 40.0%; Pred. No. 13;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 HLHNHVKKEHNAH 15
|:::|::|::|
Db      31 HVLDDGHKHKHGHGH 45
```

```
RESULT 12
US-09-612-126-1
/ Sequence 1, Application US/09612126
/ Patent No. 6284726
```

```
/ GENERAL INFORMATION:
/ APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
/ TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
/ TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAIN 5
/ FILE REFERENCE: 6056-258 CT1
/ CURRENT APPLICATION NUMBER: US/09/612,126
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/107,844
/ PRIOR FILING DATE: 1998-11-10
/ PRIOR FILING DATE: 1999-11-09
/ PRIOR FILING DATE: 1999-11-09
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 255
/ TYPE: PRT
/ ORGANISM: Human
US-09-612-126-1
```

```
Query Match          48.4%; Score 46; DB 2; Length 255;
Best Local Similarity 40.0%; Pred. No. 19;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 HLHNHVKKEHNAH 15
|:::|::|::|
Db      100 HVLDDGHKHKHGHGH 114
```

```
RESULT 13
US-09-949-002-475
/ Sequence 475, Application US/09949002
/ Patent No. 6900016
```

```
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASES, METHODS OF DETECTION
/ TITLE OF INVENTION: AND USES THEREOF
/ FILE REFERENCE: CLO00790
/ CURRENT APPLICATION NUMBER: US/09/949,002
/ PRIOR FILING DATE: 2000-01-28
/ PRIOR APPLICATION NUMBER: 60/231,401
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 10823
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 475
/ LENGTH: 579
/ TYPE: PRT
/ ORGANISM: Human
US-09-949-002-475
```

Query Match 48.4%; Score 46; DB 2; Length 579;  
Best Local Similarity 40.0%; Pred. No. 44;  
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 HLHNHVKKEHNAH 15  
DB 424 HVLHGHKHKHGHGH 438

RESULT 14  
US-09-949-002-481  
; Sequence 481, Application US/09949002  
; Patent No. 6900016  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION  
; FILE REFERENCE: CL000790  
; CURRENT APPLICATION NUMBER: US/09/949,002  
; CURRENT FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/231,401  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 10823  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 481  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-002-481

Query Match 48.4%; Score 46; DB 2; Length 579;  
Best Local Similarity 40.0%; Pred. No. 44;  
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 HLHNHVKKEHNAH 15  
DB 424 HVLHGHKHKHGHGH 438

RESULT 15  
US-09-721-154-14  
; Sequence 14, Application US/09721154  
; Patent No. 6651008  
; GENERAL INFORMATION:  
; APPLICANT: Valsberg, Eugeni  
; APPLICANT: Adams, Cynthia  
; APPLICANT: Sabry, James  
; APPLICANT: Crompton, Anne  
; TITLE OF INVENTION: Database system including computer code  
; TITLE OF INVENTION: for predictive cellular bioinformatics  
; FILE REFERENCE: CYTO007C2  
; CURRENT APPLICATION NUMBER: US/09/721,154  
; CURRENT FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: 09/311,996  
; PRIOR FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Pseudo-sequence  
US-09-721-154-14

Query Match 47.4%; Score 45; DB 2; Length 25;  
Best Local Similarity 46.7%; Pred. No. 2.2;  
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 HLHNHVKKEHNAH 15  
DB 7 HVLHGHKHKHGHGH 21

RESULT 16  
US-09-328-352-4930  
; Sequence 4930, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 4930  
; LENGTH: 363  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-4930

Query Match 47.4%; Score 45; DB 2; Length 363;  
Best Local Similarity 43.8%; Pred. No. 38;  
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 HLHNHVKKEHNAH 16  
DB 195 HHHHHHHHGHGHGH 210

RESULT 17  
US-09-540-236-2996  
; Sequence 2996, Application US/09540236  
; Patent No. 6673910  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATA  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2005-001  
; CURRENT APPLICATION NUMBER: US/09/540,236  
; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 3840  
; SEQ ID NO 2996  
; LENGTH: 367  
; TYPE: PRT  
; ORGANISM: M. catarrhalis  
US-09-540-236-2996

Query Match 47.4%; Score 45; DB 2; Length 367;  
Best Local Similarity 53.8%; Pred. No. 38;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 HNVKKEHNAH 16  
DB 321 HDDHHEHNAH 333

RESULT 18  
US-09-489-039A-9286  
; Sequence 9286, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 9286  
; LENGTH: 501  
; TYPE: PRT

ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-9286

Query Match 47.4%; Score 45; DB 2; Length 501;  
Best Local Similarity 43.8%; Pred. No. 53;  
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 HLHNVKESHAHNN 16  
| : : : : :  
DB 70 HALADIAPGSHVHNN 85

## RESULT 19

US-09-540-236-2019  
Sequence 2019, Application US/09540236  
Patent No. 6673910  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS  
FILE REFERENCE: 2709.2005-001  
CURRENT APPLICATION NUMBER: US/09/540,236  
CURRENT FILING DATE: 2000-04-04  
NUMBER OF SEQ ID NOS: 3840  
SEQ ID NO 2019  
LENGTH: 289  
TYPE: PRT  
ORGANISM: M.catarrrhalis  
US-09-540-236-2019

Query Match 46.8%; Score 44.5; DB 2; Length 289;  
Best Local Similarity 50.0%; Pred. No. 35;  
Matches 8; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 HLHNVKESHAHNN 16  
| : : : : :  
DB 229 HKLHGVA-AEHGHHH 243

## RESULT 20

US-09-991-181-95  
Sequence 95, Application US/09991181  
Patent No. 6913919  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Deanoys, Luc  
APPLICANT: Baton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerlt, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Thomas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
Acids Encoding the Same  
FILE REFERENCE: P27301C53  
CURRENT APPLICATION NUMBER: US/09/991,181  
CURRENT FILING DATE: 2001-11-16

PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105

;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089440  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089512  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089514  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089538  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089598  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089600  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089653  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089801  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089908  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089948  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089952  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090246  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090252  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25

;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 46.3%; Score 44; DB 2; Length 115;  
Best Local Similarity 37.5%; Pred. No. 16;  
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 HLHNVKKEHAHVN 16  
Db 95 HTPHHHHHHPRHH 110

RESULT 21  
US-09-990-444-95  
; Sequence 95, Application us/09990444  
; Patent No. 6930170  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Guiney, Austin L.  
; APPLICANT: Kijavlin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730P1C19  
; CURRENT APPLICATION NUMBER: US/09/990,444  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311

[illegible]





;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089600  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089653  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089801  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089908  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089948  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089952  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090246  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090252  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02

;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 46.3%; Score 44; DB 2; Length 115;  
Best Local Similarity 37.5%; Pred. No. 16;  
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 HLIHNVKKEHAAHN 16  
Db 95 HTPHHHHHHPRHH 110

RESULT 23  
US-09-992-598-95  
Sequence 95, Application US/09992598  
Patent No. 6956108  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Grimaldi, Paul J.  
APPLICANT: Guiney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730PIC20  
CURRENT APPLICATION NUMBER: US/09/992,598  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607

1	PRIOR FILING DATE: 1998-06-02	1	PRIOR APPLICATION NUMBER: 60/087609
2	PRIOR FILING DATE: 1998-06-02	2	PRIOR FILING DATE: 1998-06-18
3	PRIOR APPLICATION NUMBER: 60/087759	3	PRIOR APPLICATION NUMBER: 60/0899080
4	PRIOR FILING DATE: 1998-06-02	4	PRIOR FILING DATE: 1998-06-18
5	PRIOR APPLICATION NUMBER: 60/087827	5	PRIOR APPLICATION NUMBER: 60/0899446
6	PRIOR FILING DATE: 1998-06-03	6	PRIOR FILING DATE: 1998-06-19
7	PRIOR APPLICATION NUMBER: 60/088021	7	PRIOR APPLICATION NUMBER: 60/0899484
8	PRIOR FILING DATE: 1998-06-04	8	PRIOR FILING DATE: 1998-06-19
9	PRIOR APPLICATION NUMBER: 60/088025	9	PRIOR APPLICATION NUMBER: 60/089952
10	PRIOR FILING DATE: 1998-06-04	10	PRIOR FILING DATE: 1998-06-19
11	PRIOR APPLICATION NUMBER: 60/088026	11	PRIOR APPLICATION NUMBER: 60/0902464
12	PRIOR FILING DATE: 1998-06-04	12	PRIOR FILING DATE: 1998-06-22
13	PRIOR APPLICATION NUMBER: 60/088028	13	PRIOR APPLICATION NUMBER: 60/0902557
14	PRIOR FILING DATE: 1998-06-04	14	PRIOR FILING DATE: 1998-06-22
15	PRIOR APPLICATION NUMBER: 60/088029	15	PRIOR APPLICATION NUMBER: 60/090254
16	PRIOR FILING DATE: 1998-06-04	16	PRIOR FILING DATE: 1998-06-22
17	PRIOR APPLICATION NUMBER: 60/088030	17	PRIOR APPLICATION NUMBER: 60/0903454
18	PRIOR FILING DATE: 1998-06-04	18	PRIOR FILING DATE: 1998-06-23
19	PRIOR APPLICATION NUMBER: 60/088033	19	PRIOR APPLICATION NUMBER: 60/0903555
20	PRIOR FILING DATE: 1998-06-04	20	PRIOR FILING DATE: 1998-06-23
21	PRIOR APPLICATION NUMBER: 60/088326	21	PRIOR APPLICATION NUMBER: 60/0904299
22	PRIOR FILING DATE: 1998-06-04	22	PRIOR FILING DATE: 1998-06-24
23	PRIOR APPLICATION NUMBER: 60/088167	23	PRIOR APPLICATION NUMBER: 60/0904355
24	PRIOR FILING DATE: 1998-06-05	24	PRIOR FILING DATE: 1998-06-24
25	PRIOR APPLICATION NUMBER: 60/088202	25	PRIOR APPLICATION NUMBER: 60/0904444
26	PRIOR FILING DATE: 1998-06-05	26	PRIOR FILING DATE: 1998-06-24
27	PRIOR APPLICATION NUMBER: 60/088212	27	PRIOR FILING DATE: 1998-06-24
28	PRIOR FILING DATE: 1998-06-05	28	PRIOR APPLICATION NUMBER: 60/0904545
29	PRIOR APPLICATION NUMBER: 60/088217	29	PRIOR FILING DATE: 1998-06-24
30	PRIOR FILING DATE: 1998-06-05	30	PRIOR APPLICATION NUMBER: 60/0904727
31	PRIOR APPLICATION NUMBER: 60/088655	31	PRIOR FILING DATE: 1998-06-24
32	PRIOR FILING DATE: 1998-06-09	32	PRIOR APPLICATION NUMBER: 60/0905355
33	PRIOR APPLICATION NUMBER: 60/088734	33	PRIOR FILING DATE: 1998-06-24
34	PRIOR FILING DATE: 1998-06-10	34	PRIOR APPLICATION NUMBER: 60/0905404
35	PRIOR APPLICATION NUMBER: 60/088738	35	PRIOR FILING DATE: 1998-06-24
36	PRIOR FILING DATE: 1998-06-10	36	PRIOR APPLICATION NUMBER: 60/0905422
37	PRIOR APPLICATION NUMBER: 60/088742	37	PRIOR FILING DATE: 1998-06-24
38	PRIOR FILING DATE: 1998-06-10	38	PRIOR APPLICATION NUMBER: 60/0905577
39	PRIOR APPLICATION NUMBER: 60/088810	39	PRIOR FILING DATE: 1998-06-24
40	PRIOR FILING DATE: 1998-06-10	40	PRIOR APPLICATION NUMBER: 60/0906766
41	PRIOR APPLICATION NUMBER: 60/088824	41	PRIOR FILING DATE: 1998-06-25
42	PRIOR FILING DATE: 1998-06-10	42	PRIOR APPLICATION NUMBER: 60/0906778
43	PRIOR APPLICATION NUMBER: 60/088826	43	PRIOR FILING DATE: 1998-06-25
44	PRIOR FILING DATE: 1998-06-10	44	PRIOR APPLICATION NUMBER: 60/0906900
45	PRIOR APPLICATION NUMBER: 60/088858	45	PRIOR FILING DATE: 1998-06-25
46	PRIOR FILING DATE: 1998-06-11	46	PRIOR APPLICATION NUMBER: 60/0906944
47	PRIOR APPLICATION NUMBER: 60/088861	47	PRIOR FILING DATE: 1998-06-25
48	PRIOR FILING DATE: 1998-06-11	48	PRIOR APPLICATION NUMBER: 60/0906959
49	PRIOR APPLICATION NUMBER: 60/088876	49	PRIOR FILING DATE: 1998-06-25
50	PRIOR FILING DATE: 1998-06-11	50	PRIOR APPLICATION NUMBER: 60/0906966
51	PRIOR APPLICATION NUMBER: 60/089105	51	PRIOR FILING DATE: 1998-06-25
52	PRIOR FILING DATE: 1998-06-12	52	PRIOR APPLICATION NUMBER: 60/0908622
53	PRIOR APPLICATION NUMBER: 60/089440	53	PRIOR FILING DATE: 1998-06-26
54	PRIOR FILING DATE: 1998-06-16	54	PRIOR APPLICATION NUMBER: 60/0908633
55	PRIOR APPLICATION NUMBER: 60/089512	55	PRIOR FILING DATE: 1998-06-26
56	PRIOR FILING DATE: 1998-06-16	56	PRIOR APPLICATION NUMBER: 60/0913360
57	PRIOR APPLICATION NUMBER: 60/089514	57	PRIOR FILING DATE: 1998-07-01
58	PRIOR FILING DATE: 1998-06-16	58	PRIOR APPLICATION NUMBER: 60/091478
59	PRIOR APPLICATION NUMBER: 60/089532	59	PRIOR FILING DATE: 1998-07-02
60	PRIOR FILING DATE: 1998-06-17	60	PRIOR APPLICATION NUMBER: 60/0915444
61	PRIOR APPLICATION NUMBER: 60/089538	61	PRIOR FILING DATE: 1998-07-01
62	PRIOR FILING DATE: 1998-06-17	62	PRIOR APPLICATION NUMBER: 60/091519
63	PRIOR APPLICATION NUMBER: 60/089598	63	PRIOR FILING DATE: 1998-07-02
64	PRIOR FILING DATE: 1998-06-17	64	PRIOR APPLICATION NUMBER: 60/091626
65	PRIOR APPLICATION NUMBER: 60/089599	65	PRIOR FILING DATE: 1998-07-02
66	PRIOR FILING DATE: 1998-06-17	66	PRIOR APPLICATION NUMBER: 60/0916333
67	PRIOR APPLICATION NUMBER: 60/089600	67	PRIOR FILING DATE: 1998-07-02
68	PRIOR FILING DATE: 1998-06-17	68	PRIOR APPLICATION NUMBER: 60/091978





```
/ TYPE: PRT
/ ORGANISM: Mus Musculus
US-09-801-876B-4

Query Match
Best Local Similarity 46.3%; Score 44; DB 2; Length 404;
Matches 8; Conservative 3; Mismatches 4; Indels 6; Gaps 1;

QY 1 HLHI-----NVKKEEHAH 15
Db 142 HLHRIKPDNILLDEHGHH 162

RESULT 34
US-10-254-869-4
/ Sequence 4, Application US/10254869
/ Patent No. 6653117
/ GENERAL INFORMATION:
/ APPLICANT: YE, Jane et al
/ TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
/ TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
/ FILE REFERENCE: CL001160DIV
/ CURRENT APPLICATION NUMBER: US/10/254,869
/ CURRENT FILING DATE: 2002-09-26
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 404
/ TYPE: PRT
/ ORGANISM: Mus Musculus
US-10-254-869-4

Query Match
Best Local Similarity 46.3%; Score 44; DB 2; Length 404;
Matches 8; Conservative 3; Mismatches 4; Indels 6; Gaps 1;

QY 1 HLHI-----NVKKEEHAH 15
Db 142 HLHRIKPDNILLDEHGHH 162

RESULT 35
US-10-667-442-4
/ Sequence 4, Application US/10667442
/ Patent No. 6821765
/ GENERAL INFORMATION:
/ APPLICANT: YE, Jane et al
/ TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
/ TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
/ FILE REFERENCE: CL001160DIV II
/ CURRENT APPLICATION NUMBER: US/10/667,442
/ CURRENT FILING DATE: 2003-09-23
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 404
/ TYPE: PRT
/ ORGANISM: Mus Musculus
US-10-667-442-4

Query Match
Best Local Similarity 46.3%; Score 44; DB 2; Length 404;
Matches 8; Conservative 3; Mismatches 4; Indels 6; Gaps 1;

QY 1 HLHI-----NVKKEEHAH 15
Db 142 HLHRIKPDNILLDEHGHH 162

RESULT 36
US-09-248-796A-16084
```

```
/ Sequence 16084, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstock et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 16084
/ LENGTH: 473
/ TYPE: PRT
/ ORGANISM: Candida albicans
/ FEATURE:
/ NAME/KEY: UNSURE
/ LOCATION: (50) (429) (449)
/ OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-16084

Query Match
Best Local Similarity 46.3%; Score 44; DB 2; Length 473;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 HLHIHVKEEHAH 16
Db 284 HVLRAHAKETHGAVN 299

RESULT 37
US-07-882-292-2
/ Sequence 2, Application US/07882292
/ Patent No. 5324638
/ GENERAL INFORMATION:
/ APPLICANT: Tao, Wufan
/ APPLICANT: Lai, Eseng
/ TITLE OF INVENTION: BRAIN TRANSCRIPTION FACTOR, NUCLEIC ACIDS
/ TITLE OF INVENTION: ENCODING SAME AND USES THEREOF
/ NUMBER OF SEQUENCES: 2
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSEE: John P. White
/ STREET: c/o Cooper and Dunham, 30 Rockefeller
/ STREET: Plaza
/ CITY: New York
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 10112
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.24
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/882,292
/ FILING DATE: 19920513
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: White, John P
/ REGISTRATION NUMBER: 28,678
/ REFERENCE/DOCKET NUMBER: 41472
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-977-9550
/ TELEFAX: 212-664-0525
/ TELEEX: 422523 COOP UI
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 480 amino acids
/ TYPE: AMINO ACID
/ TOPOLOGY: linear
```



Search completed: February 11, 2006, 13:16:37  
Job time : 52 secs